

PR 19-DEC-1997; 97US-0068169.
PR 19-DEC-1997; 97US-0068365.
PR 19-DEC-1997; 97US-0068367.
PR 19-DEC-1997; 97US-0068368.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Carter KC, Duan RD, Feng P, Ferrie AM, Florence C;
PI Florence K, Greene JM, Janat F, Kyaw H, Moore PA;
PI Ni J, Rosen CA, Ruben SM, Shi Y, Soppet DR, Wei Y;
PI Yu G;
XX
XX WPI; 1999-418749/35.
DR P-PSDB; AAY36337.
XX
XX New isolated human genes encoding secreted polypeptides
PT
XX
PS Claim 1; Page 340-341; 537pp; English.

XX AAX97916 to AAX98029 represent 110 isolated human secreted protein
CC genes. AAY36224 to AAY36727 represent the secreted proteins encoded by
CC the 110 human genes. The genes and their corresponding secreted
CC polypeptides are useful for preventing, treating or ameliorating medical
CC conditions, e.g. by protein or gene therapy. Also pathological conditions
CC can be diagnosed by determining the amount of the new polypeptides in a
CC sample or by determining the presence of mutations in the new genes.
CC Specific uses are described for each of the 110 genes, based on which
CC tissues they are most highly expressed in, and include developing
CC products for the diagnosis or treatment of cancer, tumours, developmental
CC abnormalities and foetal deficiencies, blood disorders, diseases of the
CC immune system, autoimmune diseases, inflammation, allergies, Alzheimer's
CC and cognitive disorders, schizophrenia, arthritis, asthma, psoriasis,
CC sepsis, skin disorders, atherosclerosis, diabetes, cardiovascular
CC disorders, kidney disorders, digestive/endocrine disorders, infections
CC and AIDS. The polypeptides are also useful for identifying their binding
CC partners. The sequences given in AAX97907 to AAX97915 and AAY36223 are
CC used in the exemplification of the present invention.

XX Sequence 1034 BP; 217 A; 319 C; 307 G; 191 T; 0 other;

Query Match 37.5%; Score 682.2; DB 20; Length 1034;
Best Local Similarity 98.8%; Pred. No. 3.9e-146;
Matches 687; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1125 gttgcagctaagagcggtgtcaccatacagcggtacaaatggcagctgctgaccag 1184
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
330 ggaaccagctaagagcggtgtcaccatacagcggtacaaatggcagctgctgaccag 389
QY 1185 tgtgtaccagcccactgagatggcctgatgcacaaagtctcgtccgaagagcttacga 1244
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
390 tgtgtaccagcccactgagatggcctgatgcacaaagtctcgtccgaagagcttacga 449
QY 1245 catcatctccacgggcccacgcacacagccaggtgatggcgagtgccaaactcgaccct 1304
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
450 catcatctccacgggcccacgcacacagccaggtgatggcgagtgccaaactcgaccct 509
QY 1305 gcgggctgaagacatgtactcggccagagccaccagcgccacacgcgcgaagagacgg 1364
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
510 gcgggctgaagacatgtactcggccagagccaccagcgccacacgcgcgaagagacgg 569
QY 1365 caagaactctcaggtcttttagaaacccctacgtgtggagactgagtcagcggtgcgagga 1424
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
570 caagaactctcaggtcttttagaaacccctacgtgtggagactgagtcagcggtgcgagga 629
QY 1425 gagcggtcgtgattggggagggccctgagacactggccccgggcaaggggactctccagg 1484
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
630 gagcggtcgtgattggggagggccctgagacactggccccgggcaaggggactctccagg 689
QY 1485 ctctctctccccctggcagccacacatgtgccccagatgtggaagggccctccctct 1544
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
690 ctctctctccccctggcagccacacatgtgccccagatgtggaagggccctccctct 749

QY 1545 ctgccagtgtttgggtgggtgtcatgggtgtcccccacccactcctcagtggttggagt 1604
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
750 ctgccagtgtttgggtgggtgtcatgggtgtcccccacccactcctcagtggttggagt 809
QY 1605 cgaggagccaaccccccagcctcctgccaggatcaccttcggcggtcacactccagccaaata 1664
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
810 cgaggagccaaccccccagcctcctgccaggatcaccttcggcggtcacactccagccaaata 869
QY 1665 gtgttctcgggtgtgtggctggcgagcgctatgtttctctggagattcctgcaacctca 1724
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
870 gtgttctcgggtgtgtggctggcgagcgctatgtttctctggagattcctgcaacctca 929
QY 1725 agagacttcccagcgctcagcgctggatcttgcctctctgtgaggaacaagggtgccta 1784
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
930 agagacttcccagcgctcagcgctggatcttgcctctctgtgaggaacaagggtgccta 989
QY 1785 ataaatacatttctgtcttttatttaactcttaaaaaa 1819
Db |||||||||||||||||||||||||||||| |||||
990 ataaatacatttctgtctttatttaaaaaa 1024

Search completed: September 22, 2002, 17:15:48
Job time: 6081 sec

CC diagnostic markers and pharmaceutical agents for various diseases of the
CC central and peripheral nervous system, as well as cardiac, urologic and
CC gastrointestinal disorders.

Sequence 1467 BP; 258 A; 459 C; 432 G; 306 T; 12 other;
XX
SQ

Query Match 72.0%; Score 1310.2; DB 21; Length 1467;
Best Local Similarity 96.9%; Pred. No. 2.2e-289;
Matches 1423; Conservative 7; Mismatches 27; Indels 12; Gaps 9;

[illegible]

CC appetite, and can act as antigens in vaccines to raise an immune response
CC to the protein or another material cross-reactive with the protein.

XX
SQ Sequence 1323 BP; 232 A; 431 C; 384 G; 276 T; 0 other;

Query Match 72.7%; Score 1323; DB 22; Length 1323;
Best Local Similarity 100.0%; Pred. No. 2.6e-292;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 atggcattccacaagaacctgtgtgatgtgctgtggactgctctcttctgttccacagg 141
|||||
DB 1 atggcattccacaagaacctgtgtgatgtgctgtggactgctctcttctgttccacagg 60
|||||
QY 142 gctggggccagggccatgcccaaccgctgcagcaaggctcaacccctgtactac 201
|||||
DB 61 gcttggggccagggccatgcccaaccgctgcagcaaggctcaacccctgtactac 120
|||||
QY 202 aacctgtgtgaccgctctggtggcggtggtcgtctgagggcctggctggggcgggc 261
|||||
DB 121 aacctgtgtgaccgctctggtggcggtggtcgtctgagggcctggctggggcgggc 180
|||||
QY 262 atgtcaaccagtgtgtctcaacatactggtggcagactcccttctgtgcaggac 321
|||||
DB 181 atgtcaaccagtgtgtctcaacatactggtggcagactcccttctgtgcaggac 240
|||||
QY 322 accaagaacggagcctgctgggagaccaggtattctctctgtgggacccctgggcctc 381
|||||
DB 241 accaagaacggagcctgctgggagaccaggtattctctctgtgggacccctgggcctc 300
|||||
QY 382 ttctgctctgtttgctgtgtgtggaagcccgacttctccacctgtgctctctcggcgc 441
|||||
DB 301 ttctgctctgtttgctgtgtgtggaagcccgacttctccacctgtgctctctcggcgc 360
|||||
QY 442 ttctctttgggttctgtctgcacatctgtcttctgtgtgcggcctcactgtttggcc 501
|||||
DB 361 ttctctttgggttctgtctgcacatctgtcttctgtgtgcggcctcactgtttggcc 420
|||||
QY 502 ctcaactctctgcccgaagaaccacggcccgggctgggtgactcttcaactgtggct 561
|||||
DB 421 ctcaactctctgcccgaagaaccacggcccgggctgggtgactcttcaactgtggct 480
|||||
QY 562 ctgtctgtacctgtgtagaggtcatcaatacacagatggctgataccacctgggt 621
|||||
DB 481 ctgtctgtacctgtgtagaggtcatcaatacacagatggctgataccacctgggt 540
|||||
QY 622 cggggcagtgaggcggcggcctcaggcaacagcagcaggcgtggcggtggcctcc 681
|||||
DB 541 cggggcagtgaggcggcggcctcaggcaacagcagcaggcgtggcggtggcctcc 600
|||||
QY 682 cctgtgccatgcgaacatggactttgtcatggcactcactacgtcatgctgctgtg 741
|||||
DB 601 cctgtgccatgcgaacatggactttgtcatggcactcactacgtcatgctgctgtg 660
|||||
QY 742 ctgggtgctcttctgggggctggcccgcctgtgtgtggcgcgtacaaagcgtggcgtaaag 801
|||||
DB 661 ctgggtgctcttctgggggctggcccgcctgtgtgtggcgcgtacaaagcgtggcgtaaag 720
|||||
QY 802 catgggtctttgtctcctcaacacagcagcactcctgtgccatattgggtgtggatc 861
|||||
DB 721 catgggtctttgtctcctcaacacagcagcactcctgtgccatattgggtgtggatc 780
|||||
QY 862 gtcatgtactctacgcaacaagcagcaacagtcaccctgggagatgacccacacgtg 921
|||||
DB 781 gtcatgtactctacgcaacaagcagcaacagtcaccctgggagatgacccacacgtg 840
|||||
QY 922 gccatgcctctgcgcgcaaatgctgtggccttctctcttctactcgtcatccccaggtc 981
|||||
DB 841 gccatgcctctgcgcgcaaatgctgtggccttctctcttctactcgtcatccccaggtc 900
|||||
QY 982 tcccaggtgaccaggtccagcccagagcaaaagctaccagggggagacatgtacccccaccgg 1041
|||||
DB 901 tcccaggtgaccaggtccagcccagagcaaaagctaccagggggagacatgtacccccaccgg 960
|||||

QY 1042 ggcgtgggctatgagaccatccctgaagagcagaagggctcagagcatgttctgtggagaac 1101
|||||
DB 961 ggcgtgggctatgagaccatccctgaagagcagaagggctcagagcatgttctgtggagaac 1020
|||||
QY 1102 aaggccctttccatgtagcagccggttgacgtctaaagagccggtgtccaccatacagcggg 1161
|||||
DB 1021 aaggccctttccatgtagcagccggttgacgtctaaagagccggtgtccaccatacagcggg 1080
|||||
QY 1162 tacaatgggcagctgtgaccagtggtaccagcccaactgagatggcctgtatgcacaaa 1221
|||||
DB 1081 tacaatgggcagctgtgaccagtggtaccagcccaactgagatggcctgtatgcacaaa 1140
|||||
QY 1222 gttccgttccgaagagcttacgacatcatctccctccagcggccaccgccaaacagccaggtg 1281
|||||
DB 1141 gttccgttccgaagagcttacgacatcatctccctccagcggccaccgccaaacagccaggtg 1200
|||||
QY 1282 atgggcagtgccaaactgcacctgcgggctgaagacatgtactcggcccaagagccaccag 1341
|||||
DB 1201 atgggcagtgccaaactgcacctgcgggctgaagacatgtactcggcccaagagccaccag 1260
|||||
QY 1342 gcggccacacccgcgaagacgcaagaactctcaggtctttagaacccctacgtgtgg 1401
|||||
DB 1261 gcggccacacccgcgaagacgcaagaactctcaggtctttagaacccctacgtgtgg 1320
|||||
QY 1402 gac 1404
|||
DB 1321 gac 1323

RESULT 14

AAZ29294
ID AAZ29294 standard; DNA; 1467 BP.

AC AAZ29294;

XX 28-FEB-2000 (first entry)

XX Polynucleotide probe used for Northern blot analysis of MGRcm DNA.

XX Metabotropic Glutamate Receptor-like protein; MGLuR; MGRcm; G-protein;
7-transmembrane; GABA receptor; diagnostic marker; pharmaceutical agent;
central and peripheral nervous system; cardiac disorder; treatment;
urologic disorder; gastrointestinal disorder; diagnosis; seizure; anoxia;
epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;
antibody; inhibitor; screening; northern blot analysis; probe; ss.

XX Synthetic.

XX WO9960121-A1.

PN 25-NOV-1999.

XX 19-MAY-1999; 99WO-IL00265.

XX 19-MAY-1998; 98US-0085973.

XX (COMP-) COMPUGEN LTD.

XX Mintz L, Savitsky K, Toporik A;

XX WPI; 2000-086596/07.

XX New metabotropic glutamate receptor-like protein, useful for, e.g.

PT diagnosing neurological diseases

XX Disclosure; Page 66-67; 84pp; English.

XX The present sequence is the polynucleotide probe used for Northern blot
analysis of the human metabotropic glutamate receptor like protein
(MGLuR-like receptor protein), MGRcm. It detects the presence of MGRcm
coding sequence in brain, liver, kidney, stomach and heart. The MGRcm
receptor protein and encoding nucleic acid sequence, are useful as

XX PD 25-NOV-1999.
XX PF 19-MAY-1999; 99WO-IL00265.
XX PR 19-MAY-1998; 98US-0085973.
XX PA (COMP-) COMPUGEN LTD.
XX PI Mintz L, Savitsky K, Toporik A;
XX WPI; 2000-086596/07.
DR P-PSDB; AAY44273.
XX
PT New metabotropic glutamate receptor-like protein, useful for, e.g.
PT diagnosing neurological diseases -
XX
PS Claim 5; Fig 1A; 84pp; English.
XX
CC The present sequence is the DNA encoding the human metabotropic
CC glutamate receptor like protein (MGLUR-like receptor protein), MGRcm. It
CC is a novel member of the metabotropic glutamate/GABA receptor class of
CC 7-transmembrane G-protein coupled receptors showing sequence similarity.
CC It is responsible for activation of intracellular G-proteins. The MGRcm
CC receptor protein and encoding nucleic acid sequence, are useful as
CC diagnostic markers and pharmaceutical agents for various diseases of the
CC central and peripheral nervous system, as well as cardiac, urologic and
CC gastrointestinal disorders. It can be used for the diagnosis and
CC treatment of neurological and neurodegenerative disorders, such as
CC seizures, epilepsy, anoxia, Alzheimer's disease, Huntington's disease,
CC Parkinson's disease, etc. The antibody specific against MGRcm has
CC diagnostic and therapeutic applications against these disease states.
CC This receptor protein can also be employed for screening of potential
CC peptide or small molecule inhibitors of the relevant receptor-ligand
CC interaction.
XX
SQ Sequence 2406 BP; 437 A; 746 C; 759 G; 461 T; 3 other;

Query Match 98.1%; Score 1783.8; DB 21; Length 2406;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1796; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 11 cctcaccagcggaaagctacagctcggctcagcctgagctgagggagcccaaccagagcctggcc 70
DB 605 cctcaccagcggaaagctacagctcggctcagcctgagctgagggagcccaaccagagcctggcc 664
QY 71 tgggagcagagatggccatcccaaaagccttggctgagctgcttgggagctcctctctcc 130
DB 665 tgggagcagagatggccatcccaaaagccttggctgagctgcttgggagctcctctctcc 724
QY 131 tgttccaggggctggccagggccatgttccaccggctgcagccaaggcctcaacc 190
DB 725 tgttccaggggctggccagggccatgttccaccggctgcagccaaggcctcaacc 784
QY 191 cctgtactacaacctgtgtgaccgctctctgggctggtgggcatgtcctgagggcctg 250
DB 785 cctgtactacaacctgtgtgaccgctctctgggctggtgggcatgtcctgagggcctg 844
QY 251 ctggggcgccattgtccacagcttggctcaccatcctcctcctggtggcagcctccct 310
DB 845 ctggggcgccattgtccacagcttggctcaccatcctcctcctggtggcagcctccct 904
QY 311 ttgtggagaccacaagaagcctgctggggagccaccaggtattctctcttgggga 370
DB 905 ttgtggagaccacaagaagcctgctggggagccaccaggtattctctcttgggga 964
QY 371 ccttgggctcttctgctcctgtgttgcctgtgtgtgagagccagcttctcaccctgtg 430
DB 965 ccttgggctcttctgctcctgtgttgcctgtgtgtgagagccagcttctcaccctgtg 1024
QY 431 cctctcggccttctcttcttgggggttctgttgcctcctctctctctcttctgtctggcctc 490

DB 1025 cctctcggcgcttctctcttgggggttctgttgcgccatctgtcttcttctgtctggcgctc 1084
QY 491 acgtcttccctcaacttccctggcccggaagaaacacacagggcccggtggctggtgatct 550
DB 1085 agctcttgcctcaacttccctggcccggaagaaacacacagggcccggtggctggtgatct 1144
QY 551 tcaactgtgctctgctgacccctggtagaggttcaatcaatacagagtggtgctgatca 610
DB 1145 tcaactgtgctctgctgacccctggtagaggttcaatcaatacagagtggtgctgatca 1204
QY 611 tcaccctggttcggggcagtgaggagggcgccctcagggaacacagcagcagcgtggg 670
DB 1205 tcaccctggttcggggcagtgaggagggcgccctcagggaacacagcagcagcgtggg 1264
QY 671 ccgtggcctcccccctgtgcacatgcgcaacatggaacttgcacatggcactcactacgtca 730
DB 1265 ccgtggcctcccccctgtgcacatgcgcaacatggaacttgcacatggcactcactacgtca 1324
QY 731 tgctgctgctgctgggtgctcttcttctggggcctggccgccttgtgtggccctacaagc 790
DB 1325 tgctgctgctgctgggtgctcttcttctggggcctggccgccttgtgtggccctacaagc 1384
QY 791 gctggcgttaagcagtggtcttcttctcctcctcaccacagcaccctccttgcacatggg 850
DB 1385 gctggcgttaagcagtggtcttcttctcctcctcaccacagcaccctccttgcacatggg 1444
QY 851 tgggtgggacgtcatgtactatcagcgcaacagcagcacaacagtcaccacatggggtg 910
DB 1445 tgggtgggacgtcatgtactatcagcgcaacagcagcacaacagtcaccacatggggtg 1504
QY 911 accccacgtggccatgcgcctcctgcgcgaatgcctggcccttgcctctcttctacgtca 970
DB 1505 accccacgtggccatgcgcctcctgcgcgaatgcctggcccttgcctctcttctacgtca 1564
QY 971 tccccaggtctcccaggtgaccagtcagccagccagcagcaagcaccagggggacatgt 1030
DB 1565 tccccaggtctcccaggtgaccagtcagccagccagcagcaagcaccagggggacatgt 1624
QY 1031 accccacccggggcgtgggctatgagaccatcctgaaagagcagaagggtcagagcatgt 1090
DB 1625 accccacccggggcgtgggctatgagaccatcctgaaagagcagaagggtcagagcatgt 1684
QY 1091 tctggagaaacaaggccttcttccatggatgagccggttcagcgttaagagccggtgtcac 1150
DB 1685 tctggagaaacaaggccttcttccatggatgagccggttcagcgttaagagccggtgtcac 1744
QY 1151 catacagcgggtacaatgggcagctgctgaccagtggtaccagcccaactgagatggccc 1210
DB 1745 catacagcgggtacaatgggcagctgctgaccagtggtaccagcccaactgagatggccc 1804
QY 1211 tgatgcacaaagtctccgtccgaaggagcttacgacatcatctcccacggccaccgcca 1270
DB 1805 tgatgcacaaagtctccgtccgaaggagcttacgacatcatctcccacggccaccgcca 1864
QY 1271 acagccaggtgagtgggcagtgcccaactcgaacctggggctgagacatgtactcggccc 1330
DB 1865 acagccaggtgagtgggcagtgcccaactcgaacctggggctgagacatgtactcggccc 1924
QY 1331 agagccacacagcggccacacccgc- cgaaagacggcaagaactctcaggtcttttagaac 1389
DB 1925 agagccacacagcggccacacccgcgagaaagacggcaagaactctcaggtcttttagaac 1984
QY 1390 ccctacgtgtgggactgagtcagcgttggcgagagagggcggtcggtatttggggagggcc 1449
DB 1985 ccctacgtgtgggactgagtcagcgttggcgagagagggcggtcggtatttggggagggcc 2044
QY 1450 ctgaggaccttggcccgggcgaaggagcttccaggctcctcctccctcggcagggcccag 1509
DB 2045 ctgaggaccttggcccgggcgaaggagcttccaggctcctcctccctcggcagggcccag 2104
QY 1510 caacatgtgccccagatgttggaaggcctcctctctctgcccagtggttgggtgggtgcat 1569
DB 2105 caacatgtgccccagatgttggaaggcctcctctctctgcccagtggttgggtgggtgcat 2164

```
QY 83 tggccatccacaagacctgtgtgactgctgggactgctctctctcttcttccacggg 142
DB 61 tggccatccacaagacctgtgtgactgctgggactgctctctcttcttccacggg 120
QY 143 cctgggcccagggccatgtccaccggctgcagccaaaggcctcaacccctgtactaca 202
DB 121 cctgggcccagggccatgtccaccggctgcagccaaaggcctcaacccctgtactaca 180
QY 203 acctgtgacccgctctgggctgtggcactgctctgagggcgtgtgctgggcccagca 262
DB 181 acctgtgacccgctctgggctgtggcactgctctgagggcgtgtgctgggcccagca 240
QY 263 ttgtccaccgcttctgtccaccatcatctggtggccagcctcccctttgtgcaggaca 322
DB 241 ttgtccaccgcttctgtccaccatcatctggtggccagcctcccctttgtgcaggaca 300
QY 323 ccaagaaacggagcctgtcgggaccaggatattcttcttcttgggaaacctgggacctct 382
DB 301 ccaagaaacggagcctgtcgggaccaggatattcttcttcttgggaaacctgggacctct 360
QY 383 tctgctctgttctgctgtgtgtgtaagcccgacctctccacctgtgctctcggcgt 442
DB 361 tctgctctgttctgctgtgtgtgtaagcccgacctctccacctgtgctctcggcgt 420
QY 443 tctcttctgggctctgttcgccatctgtcttcttctgtctggcgctcaagcttcttgccc 502
DB 421 tctcttctgggctctgttcgccatctgtcttcttctgtctggcgctcaagcttcttgccc 480
QY 503 tcaacttctgcccgggaagaaacacgggcccgggctgggtgatatcttcaactgtggctc 562
DB 481 tcaacttctgcccgggaagaaacacgggcccgggctgggtgatatcttcaactgtggctc 540
QY 563 tgctgtgacctgtgtagaggtcatcatcaatacagagtggctgatacatcaacctggttc 622
DB 541 tgctgtgacctgtgtagaggtcatcatcaatacagagtggctgatacatcaacctggttc 600
QY 623 ggggcagtgggcaggcggccctcagggcaacagcagcagcaggctggcgctggcctccc 682
DB 601 ggggcagtgaggcggccctcagggcaacagcagcagcaggctggcgctggcctccc 660
QY 683 cctgtgccatgcgcaacatggaactttgtcatggcactcatctacgtcatgctgctgctgc 742
DB 661 cctgtgccatgcgcaacatggaactttgtcatggcactcatctacgtcatgctgctgctgc 720
QY 743 tgggtgctctctggggcctgtggccgcccctgtgtggccgctacaaagcgtggtgaagc 802
DB 721 tgggtgctctctggggcctgtggccgcccctgtgtggccgctacaaagcgtggtgaagc 780
QY 803 atggggtcttctgtctctcaccacagccacctccgttgccatatggtgggtggtgatacg 862
DB 781 atggggtcttctgtctctcaccacagccacctccgttgccatatggtgggtggtgatacg 840
QY 863 tcatgtatacttacggcaacagcagcaacacagtcctccacctgggatatgccccacgctgg 922
DB 841 tcatgtatacttacggcaacagcagcaacacagtcctccacctgggatatgccccacgctgg 900
QY 923 ccacgcccctcggccaatgctgggcttgcctctcttcttctacgtcatcccgaggtct 982
DB 901 ccacgcccctcggccaatgctgggcttgcctctcttcttctacgtcatcccgaggtct 960
QY 983 ccaggtgaccaggtccagccagagcaaaagctacaggggggacatgtacccccaccggg 1042
DB 961 ccaggtgaccaggtccagccagagcaaaagctacaggggggacatgtacccccaccggg 1020
QY 1043 gcgtgggctatgagaccatctctgaagagcagaaggggtcagagcatgttctgtggaaca 1102
DB 1021 gcgtgggctatgagaccatctctgaagagcagaaggggtcagagcatgttctgtggaaca 1080
QY 1103 aggcctttccatgtagagccggttgagctaaaggccggtgtaccatacagcgggt 1162
DB 1081 aggcctttccatgtagagccggttgagctaaaggccggtgtaccatacagcgggt 1140
QY 1163 acaatgggcagctgctgaccagtggtaccagccacctgagatggccctgtagcacaaag 1222
```

```
DB 1141 acaatgggcagctgctgaccagtggtaccagccacctgagatggccctgtagcacaaag 1200
QY 1223 ttccgtccgaagagetttaagacatcatctccacggggccaccgcaaacagccaggtga 1282
DB 1201 ttccgtccgaagagetttaagacatcatctccacggggccaccgcaaacagccaggtga 1260
QY 1283 ttggcagtgccaaactgaccctgggctgaagacatgtactcggccagagccaccagg 1342
DB 1261 ttggcagtgccaaactgaccctgggctgaagacatgtactcggccagagccaccagg 1320
QY 1343 cggccacacgcgcgaagacggcaagaactctcaggtcttttagaaacccctacgttggg 1402
DB 1321 cggccacacgcgcgaagacggcaagaactctcaggtcttttagaaacccctacgttggg 1380
QY 1403 actgagtcagcggctggcaggagagcggctcgagatttgggaggccctcagagacctggc 1462
DB 1381 actgagtcagcggctggcaggagagcggctcgagatttgggaggccctcagagacctggc 1440
QY 1463 cccgggcaaggagactctccaggtctctctccctccctcggcagcccgcaacatgtgccc 1522
DB 1441 cccgggcaaggagactctccaggtctctctccctccctcggcagcccgcaacatgtgccc 1500
QY 1523 agatgtggaaggccctccctctctcagtggttgggtgggtgtcatgggtgtcccccacc 1582
DB 1501 agatgtggaaggccctccctctctcagtggttgggtgggtgtcatgggtgtcccccacc 1560
QY 1583 cactctcagtggttgtgagtcgagagagcgaacccacgacctcctcgcagagatcaactcg 1642
DB 1561 cactctcagtggttgtgagtcgagagagcgaacccacgacctcctcgcagagatcaactcg 1620
QY 1643 gcggtcacactccagccaaatagttcttcggggtgtgtggctggcagcgcctatgttc 1702
DB 1621 gcggtcacactccagccaaatagttcttcggggtgtgtggctggcagcgcctatgttc 1680
QY 1703 tctggagattctctgcaacctcaagagactcccgagcgcctcagcgcctggatcttgcct 1762
DB 1681 tctggagattctctgcaacctcaagagactcccgagcgcctcagcgcctggatcttgcct 1740
QY 1763 ctgtgaggaaacagggtgcctataataacattctgctttataactcttaa 1816
DB 1741 ctgtgaggaaacagggtgcctataataacattctgctttataataaaaaaa 1794

RESULT 10
AAZ29292
ID AAZ29292 standard; DNA; 2406 BP.
XX
AC AAZ29292;
XX
DT 28-FEB-2000 (first entry)
XX
DE Human Metabotropic Glutamate Receptor-like protein, MGRcm encoding DNA.
XX
KW Metabotropic Glutamate Receptor-like protein; MGLuR; MGRcm; G-protein;
KW 7-transmembrane; GABA receptor; diagnostic marker; pharmaceutical agent;
KW central and peripheral nervous system; cardiac disorder; treatment;
KW urologic disorder; gastrointestinal disorder; diagnosis; anoxia;
KW epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;
KW antibody; inhibitor; screening; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 541..1980
FT /*tag= a
FT /product= "MGRcm protein"
FT /note= "Metabotropic Glutamate Receptor-like protein"
FT mat_peptide 601..1977
FT /*tag= b
FT /label= MGRcm_mature_protein
XX
PN W09960121-A1.
```


	Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
	WPI; 2001-093989/11.
	P-PSDB; AAB88359.
	Nucleic acids encoding secretory proteins/membrane proteins, useful in
	gene therapy or as candidate target molecules in drug development -
	Claim 1; SEQ ID 85; 609pp + CD ROM; English.
	This invention relates to nucleic acid sequences AAF93744 - AAF93916
	which encode human secretory or membrane proteins represented by
	AAB88317 - AAB88419. Included in the invention are primers
	AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
	cDNA sequences of the invention. The invention also includes methods for
	the production of antibodies directed against the proteins, and cDNA
	sequences, which can be used in vaccines. The polynucleotide sequences
	can be used in gene therapy. The polynucleotide sequences and the
	proteins they encode may be used in the prevention, treatment and
	diagnosis of diseases associated with inappropriate secretory
	protein/membrane protein expression. The nucleic acids and complementary
	sequences may also be used as DNA probes in diagnostic assays
	(e.g. polymerase chain reactions (PCR)) to detect and quantitate the
	presence of similar nucleic acid sequences in samples. They may also be
	used to study the expression and function of secretory proteins/membrane
	polypeptides and their role in metabolism. The polypeptides may be used
	as antigens in the production of antibodies against them and in assays to
	identify modulators (agonists and antagonists) of expression and
	activity. The antibodies and antagonists may also be used as therapeutic
	agents to down regulate expression and activity. The antibodies may also
	be used as diagnostic agents for detecting the presence of the
	polypeptides in samples (e.g. by enzyme linked immunosorbant assay
	(ELISA). Examples of diseases which may be treated include rheumatoid
	arthritis and diabetes.
	Sequence 1808 BP; 321 A; 580 C; 531 G; 376 T; 0 other;
	Query Match 98.1%; Score 1785; DB 22; Length 1808;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 1796; Conservative 0; Mismatches 0; Indels 1; Gaps 1
QY	10 ccctcaccagccgaaagtacgactcggtcgcagctcgaggagaccacacgagcctggc 69
Db	
QY	13 ccctcaccagccgaaagtacgactcggtcgcagctcgaggagaccacacgagcctggc 72
Db	
QY	70 ctgggagccaggatgccatccaagaacctggtagtgccttggaactcctctcttc 129
Db	
QY	73 ctgggagccaggatgccatccaagaacctggtagtgccttggaactcctctcttc 132
Db	
QY	130 ctgttcccaggcctgggcccaggcccatgtccccaccgctgcagcaaggcctcaac 189
Db	
QY	133 ctgttcccaggcctgggcccaggcccatgtccccaccgctgcagcaaggcctcaac 192
Db	
QY	190 cccctgtactacaacctgtgtgaccgctctgggcgctgggcatcgctcggagccggtg 249
Db	
QY	193 cccctgtactacaacctgtgtgaccgctctgggcgctgggcatcgctcggagccggtg 252
Db	
QY	250 gctggggcgggcattgtcacccagtttgtctcacatactcctggtggccagcctcccc 309
Db	
QY	253 gctggggcgggcattgtcacccagtttgtctcacatactcctggtggccagcctcccc 312
Db	
QY	310 ttgtgtcagacaccaaagaaacgagcctgctggggaccaggtattcttccttgggg 369
Db	
QY	313 ttgtgtcagacaccaaagaaacgagcctgctggggaccaggtattcttccttgggg 372
Db	
QY	370 accctgggcctctctgcctcgttgtttgctgtgtgtgaagcccgactcttccacctgt 429
Db	

PS The invention relates to polynucleotides encoding novel human
XX proteins or their active domains. The polypeptides, polynucleotides and
CC antibodies raised against the polypeptides are used in a method of
CC treatment of a mammal and prevention of disorders caused by the aberrant
CC protein expression or activity. The polypeptides can be used as
CC molecular weight markers, food supplements, and in antibody production.
CC The polypeptides are used to identify compounds which bind to the
CC polypeptides. Polynucleotides of the invention are used as probes and
CC primers, for sequencing, for chromosome or gene mapping, in the
CC production of recombinant proteins, and in generating anti-sense DNA or
CC RNA and in gene therapy. Polypeptides of the invention can be used to
CC target drugs to a tumour, in assays to determine biological activity, to
CC raise antibodies/elicit an immune response, to determine quantitative
CC protein levels, as tissue markers, and to isolate receptors or ligands.
CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue, wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma,
CC graft-versus-host disease, eczema, haemophilia, thrombosis,
CC anti-inflammatory diseases, nervous system disorders, and infection.
XX The present sequence encodes a protein of the invention.
SQ Sequence 2027 BP; 370 A; 653 C; 607 G; 397 T; 0 other;

Query Match 98.8%; Score 1797.4; DB 22; Length 2027;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1801; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 10 cctcaccagcgaagtacgagtcgctcagctcagctcggaggaccacacagagcctgac 69
DB 221 cctcaccagcgaagtacgagtcgctcagctcggaggaccacacagagcctgac 280
QY 70 ctggggagcagagtgccatccacaaagccttgggtgatgctgctggagctgctcttc 129
DB 281 ctggggagcagagtgccatccacaaagccttgggtgatgctgctggagctgctcttc 340
QY 130 ctgttccacagggcctggggcagagccatgtccaccgctgagccagcagccacac 189
DB 341 ctgttccacagggcctggggcagagccatgtccaccgctgagccagcagccacac 400
QY 190 cccctgtactacaaacctgtgtgaccgctctggggcgtggggcagctcctggagccgtg 249
DB 401 cccctgtactacaaacctgtgtgaccgctctggggcgtggggcagctcctggagccgtg 460
QY 250 gctggggcgggcatgttcacacagcttggctcaccatcatcctggtggcagcctccc 309
DB 461 gctggggcgggcatgttcacacagcttggctcaccatcatcctggtggcagcctccc 520
QY 310 ttgtgcagacacacaaagagagcctgctgggagccaggtattctcctctgggg 369
DB 521 ttgtgcagacacacaaagagagcctgctgggagccaggtattctcctctgggg 580
QY 370 accctgggctctctctgctgctgtttgctgtgtgtggaagcccgactctccactgt 429
DB 581 accctgggctctctctgctgctgtttgctgtgtgtggaagcccgactctccactgt 640
QY 430 gctctcgggctctctctggtggttctgtgctcactgtctctctctctgagcgct 489
DB 641 gctctcgggctctctctggtggttctgtgctcactgtctctctctctgagcgct 700
QY 490 cagctcttgcctcacttctggcccggaagaaacacagggcccgccgggctgggtgac 549
DB 701 cagctcttgcctcacttctggcccggaagaaacacagggcccgccgggctgggtgac 760
QY 550 ttactgtggtctgctgctgacccctggtagaggtcatcatcaatcagagtggtgctgac 609

DB 761 ttactgtggtctgctgctgacccctggtagaggtcatcatcaatcagagtggtgctgac 820
QY 610 ataccctgttccggggcagtgccgagggcgccctcagggcacaacagcagcagcgtcg 669
DB 821 ataccctgttccggggcagtgccgagggcgccctcagggcacaacagcagcagcgtcg 880
QY 670 gctgtggtccctccctctgctcgcacacacatgagcttctgctggtgacacacacac 729
DB 881 gctgtggtccctccctctgctcgcacacacatgagcttctgctggtgacacacacac 940
QY 730 atgtgtgtgctgctggtgctcctctctggtggggcctggcccgccctgtgtggcctacaag 789
DB 941 atgtgtgtgctgctggtgctcctctctggtggggcctggcccgccctgtgtggcctacaag 1000
QY 790 cgtgtggtgaagcagtggtgtgtgtgtgtctctcaccacagcagcagcctcgttgccatg 849
DB 1001 cgtgtggtgaagcagtggtgtgtgtgtgtgtctctcaccacagcagcagcctcgttgccatg 1060
QY 850 gtgtgtgtgctgctgctgctcctctctggtggggcctggcccgccctgtgtggcctggat 909
DB 1061 gtgtgtgtgctgctgctgctcctctctggtggggcctggcccgccctgtgtggcctggat 1120
QY 910 gacccacgctggcagcagcctcctcgcgcgcaatgcctggcctcgtcctctctctcagtc 969
DB 1121 gacccacgctggcagcagcctcctcgcgcgcaatgcctggcctcgtcctctctcagtc 1180
QY 970 atccccaggtctccaggtgacaaagtcacagccagcagcagcagcagcagcagcagcag 1029
DB 1181 atccccaggtctccaggtgacaaagtcacagccagcagcagcagcagcagcagcagcag 1240
QY 1030 taccacacccggcggtgggtatgagaccatcctgaaagagcaggaagggctcagagcagtg 1089
DB 1241 taccacacccggcggtgggtatgagaccatcctgaaagagcaggaagggctcagagcagtg 1300
QY 1090 ttctgtggaacaaagccttttccatgagcagcgttgagcagtaagagccggtgtgca 1149
DB 1301 ttctgtggaacaaagccttttccatgagcagcgttgagcagtaagagccggtgtgca 1360
QY 1150 ccatacagcgggtacaatgggagcgtgtgaccagctgtgtaccagcccaactgagatggcc 1209
DB 1361 ccatacagcgggtacaatgggagcgtgtgaccagctgtgtaccagcccaactgagatggcc 1420
QY 1210 ctgtatgcacaaagtctcgtccgaagagcctcagcagcagcagcagcagcagcagcagc 1269
DB 1421 ctgtatgcacaaagtctcgtccgaagagcctcagcagcagcagcagcagcagcagcagc 1480
QY 1270 aacagcagctgagtgagcagtgccactcgcctgagcagcagcagcagcagcagcagcagc 1329
DB 1481 aacagcagctgagtgagcagtgccactcgcctgagcagcagcagcagcagcagcagcagc 1540
QY 1330 cagagccacagcggcggccacacgcgcggaagacggaagacggaagacggaagacggaagac 1389
DB 1541 cagagccacagcggcggccacacgcgcggaagacggaagacggaagacggaagacggaagac 1600
QY 1390 cctcagctgtggagcagtgagcagtgccagcagcagcagcagcagcagcagcagcagcagc 1449
DB 1601 cctcagctgtggagcagtgagcagtgccagcagcagcagcagcagcagcagcagcagcagc 1660
QY 1450 ctgagcagctggcccgcgccaggggactctccagcagcagcagcagcagcagcagcagcagc 1509
DB 1661 ctgagcagctggcccgcgccaggggactctccagcagcagcagcagcagcagcagcagcagc 1720
QY 1510 caacatgtgcccagatgtggaagggcctcctctctctcagtgagcagtggtgtggtggtggtat 1569
DB 1721 caacatgtgcccagatgtggaagggcctcctctctctcagtgagcagtggtgtggtggtggtat 1780
QY 1570 ggggtgtcccaac 1629
DB 1781 ggggtgtcccaac 1840
QY 1630 caggatcacctcgggctgac 1689
DB 1841 caggatcacctcgggctgac 1900

[illegible]

D	b	1335	cagagccaccaggcgccacacccgccaagagcgcgcgaagaacctctcaggctctttgagaac	1394
Q	y	1390	ccctaactgtgggactgagtcagcgttgccgagagagagcgcgtcggatttggggagggcc	1449
D	b	1395	ccctaactgtgggactgagtcagcgttgccgagagagcgcgtcggatttggggagggcc	1454
Q	y	1450	ctgaggacctggccccggcgaaggagactctccagctcctcctccccctggcaggcccg	1509
D	b	1455	ctgaggacctggccccggcgaaggagactctccagctcctcctccccctggcaggcccg	1514
Q	y	1510	caacatgtgccccagatgtgggaaggccctccctctctcgccagtggttgggtgggtgcat	1569
D	b	1515	caacatgtgccccagatgtgggaaggccctccctctctcgccagtggttgggtgggtgcat	1574
Q	y	1570	gggtgtccccaccactcctcagtggtttgtgagtcgagagcgaacccccagcctcctgc	1629
D	b	1575	gggtgtccccaccactcctcagtggtttgtgagtcgagagcgaacccccagcctcctgc	1634
Q	y	1630	caggatcaactgcgcggtcacactccagccaaatagttctcggggtgggtgggtggcca	1689
D	b	1635	caggatcaactgcgcggtcacactccagccaaatagttctcggggtgggtgggtggcca	1694
Q	y	1690	gcgcctatgtttctctggagattcctgcgaacctcgaagagacttcccagcgcgtcaggcct	1749
D	b	1695	gcgcctatgtttctctggagattcctgcgaacctcgaagagacttcccagcgcgtcaggcct	1754
Q	y	1750	ggaattgtcctcctgtggaacaaagggtgcctaataatattcttcttattaaac	1809
D	b	1755	ggaattgtcctcctgtggaacaaagggtgcctaataatattcttcttattaaac	1814
Q	y	1810	tccttaaaaaa	1819
D	b	1815	aaaaaaaaa	1824
RESULT 4				
AAS22707				
ID	AAS22707 standard; cDNA; 2027 BP.			
XX	AAS22707;			
XX	24-OCT-2001 (first entry)			
DT	Human cDNA encoding a novel human protein #273.			
DE	Human; novel protein; ss; Antianaemic; osteopathic; antiinflammatory;			
KW	immunomodulatory; cystostatic; neuroprotective; vulnerrary; nootropic;			
KW	anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;			
KW	antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;			
KW	thrombolytic; immunogen; antibody; gene therapy; neurological disorder;			
KW	Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;			
KW	tissue regeneration; immune disorder.			
XX	Homo sapiens.			
OS	Homo sapiens.			
XX	WO200155437-A2.			
PN	02-AUG-2001.			
PD	25-JAN-2001; 2001WO-US02623.			
XX	25-JAN-2000; 2000US-0491404.			
PF	(HYSE-) HYSEQ INC.			
XX	Tang YT, Liu C, Drmanac RT;			
XX	WPI: 2001-451939/48.			
XX	P-PSDB; AAU14402.			
DR	Isolated polypeptides useful for treating anti-inflammatory diseases,			
XX	nervous system disorders, and for regenerating bone and cartilage -			
PT				
XX				

Qy	1741	ctcaggcctgattctgctctctctgtagagaaacagggtgcctaataatacatattctgc	1800
Db	1741	ctcaggcctgattctgctctctctgtagagaaacagggtgcctaataatacatattctgc	1800
Qy	1801	tttattaactcttaaaaaa	1819
Db	1801	tttattaactcttaaaaaa	1819
RESULT	2		
AAS22471			
ID	AAS22471	standard; cDNA; 2084 BP.	
XX	AAS22471;		
AC			
XX	24-OCT-2001	(first entry)	
DT			
XX	Human cDNA encoding a novel human protein #37.		
DE			
XX	Human; novel protein; ss; Antianaemic; osteopathic; antiinflammatory;		
KW	immunomodulatory; cytostatic; neuroprotective; vulnary; nootropic;		
KW	anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;		
KW	antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;		
KW	thrombolytic; immunogen; antibody; gene therapy; neurological disorder;		
KW	Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;		
KW	tissue regeneration; immune disorder.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200155437-A2.		
XX			
PD	02-AUG-2001.		
XX			
PF	25-JAN-2001; 2001WO-US02623.		
XX			
PR	25-JAN-2000; 2000US-0491404.		
XX	(HYSE-) HYSEQ INC.		
PA			
XX	Tang YT, Liu C, Drmanac RT;		
PI			
XX	WPI; 2001-451939/48.		
DR	P-PSDB; AAU14166.		
DR			
XX			
PT	Isolated polypeptides useful for treating anti-inflammatory diseases,		
PT	nervous system disorders, and for regenerating bone and cartilage -		
XX			
PS	Claim 1; Page 226-228; 894pp; English.		
XX			
CC	The invention relates to polynucleotides encoding novel human		
CC	proteins or their active domains. The polypeptides, polynucleotides and		
CC	antibodies raised against the polypeptides are used in a method of		
CC	treatment of a mammal and prevention of disorders caused by the aberrant		
CC	protein expression or activity. The polypeptides can be used as		
CC	molecular weight markers, food supplements, and in antibody production.		
CC	The polypeptides are used to identify compounds which bind to the		
CC	polypeptides. Polynucleotides of the invention are used as probes and		
CC	primers, for sequencing, for chromosome or gene mapping, in the		
CC	production of recombinant proteins, and in generating anti-sense DNA or		
CC	RNA and in gene therapy. Polypeptides of the invention can be used to		
CC	target drugs to a tumour, in assays to determine biological activity, to		
CC	raise antibodies/elicit an immune response, to determine quantitative		
CC	protein levels, as tissue markers, and to isolate receptors or ligands.		
CC	Polypeptides of the invention may also be useful in treating platelet		
CC	disorders, stem cell disorders, regenerating bone, cartilage, tendon,		
CC	ligament and/or nerve tissue, wound healing, treating burns, promoting		
CC	the proliferation, differentiation and survival of stem cells, as a		
CC	contraceptive, treating osteoporosis and osteoarthritis, anaemia,		
CC	Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral		
CC	sclerosis, stroke, immune deficiencies resulting from bacterial, viral or		
CC	clonal infection or from autoimmunity, cancer, allergy, asthma,		
CC	graft-versus-host disease, eczema, haemophilia, thrombosis		
CC	anti-inflammatory diseases, nervous system disorders, and infection.		

DR P-PSDB: AAY57283.

Human G protein coupled protein receptor peptides useful for the prevention, diagnosis and treatment of cell proliferative, neurological and immune disorders -

PS Claim 9: Page 67; 71pp; English.

The invention provides human G protein coupled protein receptor (HGPRP) polypeptides and polynucleotides encoding them. The polypeptides can be produced by standard recombinant methodology. The polynucleotides and polypeptides may be used in the prevention, treatment and diagnosis of diseases associated with their inappropriate expression. Diseases that can be treated are cell proliferative disorders (e.g. arteriosclerosis, atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease and Parkinson's disease). The anti-HGPRP antibodies may also be used as diagnostic agents for detecting the presence of HGPRP polypeptides in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). Sequences AAZ90521-526 represent cDNA fragments encoding the HGPRP polypeptides (AAZ57283-288).

Sequence 1824 BP; 331 A; 581 C; 535 G; 377 T; 0 other;

Query Match	100.0%;	Score 1819;	DB 21;	Length 1824;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1819: Conservative	0;	Mismatches	0;	Indels 0;

Qy	1	cggctcgagccctaccagccggaagtaacgctcgagctggaggaccaca	60
Db	1	cggctcgagccctaccagccggaagtaacgctcgagctggaggaccaca	60
Qy	61	gagcctggcctggagccagatggcatccacaagccttgatgtgacctggactg	120
Db	61	gagcctggcctggagccagatggcatccacaagccttgatgtgacctggactg	120
Qy	121	ccctctctctgttccagggcctggcccagggcctgtcccaccggctgcagccaa	180
Db	121	ccctctctctgttccagggcctggcccagggcctgtcccaccggctgcagccaa	180
Qy	181	ggcctaaacccctgtactacaacctgtgtgacgcgtcttgggcctggggcatcgtcctg	240
Db	181	ggcctaaacccctgtactacaacctgtgtgacgcgtcttgggcctggggcatcgtcctg	240
Qy	241	gagccctggctgggcgggcattgtcacacgtttgtcaccatcatcctgtggcc	300
Db	241	gagccctggctgggcgggcattgtcacacgtttgtcaccatcatcctgtggcc	300
Qy	301	agctccccccttgtcaggacaccaaagacggagcctgtgggaccagattattctc	360
Db	301	agctccccccttgtcaggacaccaaagacggagcctgtgggaccagattattctc	360
Qy	361	ctctcgggaccctggcctctctcctccctgtgttgccctgtgtgtagagccgaattc	420
Db	361	ctctcgggaccctggcctctctcctccctgtgttgccctgtgtgtagagccgaattc	420
Qy	421	tccacctgtgcctctcgccctctcctctttgggggtctgttcgcacatcgtctcttgt	480
Db	421	tccacctgtgcctctcgccctctcctctttgggggtctgttcgcacatcgtctcttgt	480
Qy	481	ctggcggctcaagtcttttgccctcaactctctggcccgaagaaacacacggcccggggc	540
Db	481	ctggcggctcaagtctttgccctcaactctctggcccgaagaaacacacggcccggggc	540
Qy	541	tgggtgatcttcaacttggtctctgtcgtcgaccctgtagaggtcatcatcaatcacag	600
Db	541	tgggtgatcttcaacttggtctctgtcgtcgaccctgtagaggtcatcatcaatcacag	600
Qy	601	tggctgatcatcacctcctgggtcgggcagtggcaggcgccctcaggggcaacagcagc	660
Db	601	tggctgatcatcacctcctgggtcgggcagtggcaggcgccctcaggggcaacagcagc	660

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 15:34:27 ; Search time 292.25 seconds
(without alignments)
10686.287 Million cell updates/sec

Title: US-09-895-686-7
Perfect score: 1819
Sequence: 1 cggtcgcagccctaccagc.....ctttatttaactcttaaaaaa 1819

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*

1:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1980.DAT:*
2:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1981.DAT:*
3:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1982.DAT:*
4:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1983.DAT:*
5:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1984.DAT:*
6:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1985.DAT:*
7:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1986.DAT:*
8:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1987.DAT:*
9:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1988.DAT:*
10:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1989.DAT:*
11:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1990.DAT:*
12:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1991.DAT:*
13:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1992.DAT:*
14:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1993.DAT:*
15:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1994.DAT:*
16:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1995.DAT:*
17:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1996.DAT:*
18:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1997.DAT:*
19:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1998.DAT:*
20:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT:*
21:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA2000.DAT:*
22:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA2001A.DAT:*
23:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA2001B.DAT:*
24:	/SIDSL/gcgdata/geneseq/geneseq-n-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match %	Length DB ID	Description
1	1819	100.0	1824	21 AAZ90521 Human GPCR protein
2	1799	98.9	2084	22 AAS22471 Human cDNA encodin
3	1798.8	98.9	1936	21 AAZ43798 Human fetal brain
4	1797.4	98.8	2027	22 AAS22707 Human cDNA encodin
5	1797	98.8	1866	22 AAF94491 Human hydrophobic
6	1786.4	98.2	1860	20 AAX97994 Human secreted pro
7	1785	98.1	1808	22 AAF93786 Human cDNA encodin
8	1785	98.1	1880	22 AAI59005 Human polynucleoti
9	1784.4	98.1	1794	22 AAI60791 Human polynucleoti

10	1783.8	98.1	2406	21 AAZ29292 Human Metabotropic
11	1630.4	89.6	1790	20 AAZ32502 G-protein coupled
12	1324.4	72.8	1326	20 AAZ32501 G-protein coupled
13	1323	72.7	1323	22 AAF94481 Human hydrophobic
14	1310.2	72.0	1467	21 AAZ29294 Polynucleotide pro
15	682.2	37.5	1034	20 AAX98029 Human secreted pro
16	640.8	35.2	684	22 AAH03345 Human cDNA clone (
17	624	34.3	891	21 AAZ29293 Mouse Metabotropic
18	587.8	32.3	946	22 AAI99248 Human excretory re
19	587.8	32.3	946	22 AAI63598 Human kidney relat
20	575.8	31.7	613	22 AAF93982 Primer specific fo
21	271.8	14.9	3362	21 AAC77294 Human ORFX ORF2849
22	270.2	14.9	1212	20 AAZ20297 Human G-protein co
23	270.2	14.9	1520	21 AAZ90525 Human GPCR protein
24	270.2	14.9	2389	22 AAH51748 Human polynucleoti
25	270.2	14.9	2749	22 AAK94479 Human full-length
26	270.2	14.9	3852	20 AAZ20298 Human G-protein co
C 27	255	14.0	540	22 AAH08883 Human cDNA clone (
C 28	228.4	12.6	251	22 AAF94129 primer specific fo
C 29	207.8	11.4	566	22 AAI99247 Human excretory re
30	207.8	11.4	566	22 AAI63597 Human kidney relat
31	202.8	11.1	207	20 AAX40620 Human secreted pro
C 32	202.4	11.1	1000	22 ABA48935 Human breast cell
C 33	202.4	11.1	1000	22 ABA66850 Human foetal liver
C 34	202.4	11.1	1000	22 ABA33918 Probe #12384 for g
C 35	202.4	11.1	1000	22 AAK15285 Human brain expres
C 36	202.4	11.1	1000	22 AAK41008 Human bone marrow
C 37	202.4	11.1	1000	22 AAI21778 Probe #11711 for g
C 38	202.4	11.1	1000	22 AAI07463 Probe #15747 used
C 39	202.4	11.1	1000	22 AAI07463 Probe #7454 used t
C 40	201.8	11.1	205	21 AAC00605 Human RECAP polynu
41	201.4	11.1	1619	22 AAF58615 Human cDNA sequenc
42	201.4	11.1	2446	22 AAH14688 Human cervical can
43	201.4	11.1	3371	22 AAH72766 Receptor #24 part1
44	191.6	10.5	2484	22 AAF80536 Extended human sec
45	187.6	10.3	1114	20 AAX97687

ALIGNMENTS

RESULT	1
AAZ90521	
ID	AAZ90521 standard; cDNA; 1824 BP.
XX	AAZ90521;
AC	AAZ90521;
XX	05-JUN-2000 (first entry)
DT	Human GPCR protein (HGPRP) encoding cDNA (clone ID 1258981).
XX	Human; G protein coupled protein receptor; HGPRP; cell proliferation;
DE	neurological; immune disorder; cytostatic; anti-arteriosclerotic;
XX	anti-atherosclerotic; hepatotropic; antiinflammatory; virucide; leukemia;
XX	immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;
XX	anti-Alzheimer's; anti-Parkinsonian; gene therapy; ss.
OS	Homo sapiens.
XX	WO200015793-A2.
PN	23-MAR-2000.
XX	
PD	
XX	17-SEP-1999; 99WO-US20958.
PF	
XX	17-SEP-1998; 98US-0156513.
XX	(INCY-) INCYTE PHARM INC.
PA	
XX	Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Gorgone GA;
PI	Baughn MR;
XX	WPI; 2000-271432/23.
DR	

FILING DATE: 04-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Reiter, Stephen E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: FP41 9772
TELEPHONE: 619-546-4737
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 906 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-367-264-2

Query Match 6.0%; Score 140; DB 3; Length 906;
Best Local Similarity 22.5%; Pred. No. 8.4e-06;
Matches 80; Conservative 50; Mismatches 140; Indels 86; Gaps 17;
QY 20 GAWAQGHVPPCCSQGLNPLYYNLCDSRGAWGIVLEAVAGAGIVTTFVLITILV--ASLPE 77
Db 568 GAW-----PNADLTGCEPIPVRYLEWSNIESIIAIAFSCILGILVTLFVTLFVLYRTPV 622
QY 78 VQDTKKRSLGTVQFFELGT-LGLFCLVFACVVKPDFSTCASRRFLEGLVFAICFSCLA 136
Db 623 VKSSREL-----YIILAGIFLGYVC-PFTLIAKPTTSCYLQRLVLGLSSAMCYSA-- 675
QY 137 HVPALNFLARKNHG-----PR---GWVFTVALLTLVEVIINTEWLIITLVRGSG 184
Db 676 -VTKTNRIARILAGSKKIKCTKPREMSAWAQVIAISILSVOLT-----LVVTLI---- 725
QY 185 EGGPQGNSSKGNVAVSP-----IMEPMPILSYPSIKYVYLICNTSNLGVAVPLGYNGLLIMSCITYAF 772
Db 726 -----PALCGRYKRWKRGHGVFLLTTATSAIWWIVMVTYGNKQHNSTWDDPTLAI 283
QY 229 -----PALCGRYKRWKRGHGVFLLTTATSAIWWIVMVTYGNKQHNSTWDDPTLAI 283
Db 773 KTRNVFANFNK-----YIAFTWYTTCTIILAFVPIY-FGSNYKIITTCFAVLSVT 824
QY 284 LAANAWAFVLFFVI---PE---VSQVTKSSPEQSYOGD-MYPTRGVGYETILKEQK 332
Db 825 VALGCMFTPKMYIIIAKPERNVRSFTTSDVVRMHVGDGKLPKCRSNTFLNIFRRKK 880

Search completed: September 16, 2002, 09:33:09
Job time: 168 sec

Qy	229	-----PALCGRYRKRKHGVFVLLLTATSVAIWVMVMTYGKNOHNSPTWDDPFLATA	283
D6	773	KTRNVPANFEAK-----YIAFTMYTTCIIWLAFVPIY-FGSNNYKIITTCFAVSLSVT	824
Qy	284	LAANAFAVLFFYVI---PE---VSQVTKSSEQSOGD-MYPTRGVGYETILKEOK-GOS	335
D6	825	VALGCMTPKMYIILIAKPENRVSAFTTSDVVRMHVGDKLPCRSNTFLNFRKKAPGAG	884
Qy	336	MIVEN-KAFSMDPEVAAKRP	354
D6	885	NANSNGKSVSPSPGGROAP	904

```

RESULT 12
US-08-465-157-2
; Sequence 2, Application US/08465157
; Patent No. 5869609
; GENERAL INFORMATION:
; APPLICANT: Mulvihill, Eileen R
; APPLICANT: Hagen, Frederick S
; APPLICANT: Houamed, Khalel M
; APPLICANT: Almers, Wolfhard
; TITLE OF INVENTION: G PROTEIN COUPLED GLUTAMATE
; TITLE OF INVENTION: RECEPTORS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94105

```

Query Match	26.1%	Score 142:	DB 2:	Length 1199:
Best Local Similarity	22.4%	Pred. No. 7.9e-06:		
Matches	85:	Conservative	54:	Mismatches 153:
				Indels 88: Gaps 19:
QY	20	GANAQGHVPCGSQLNPLYNLCDRSGANGIVLEAVAGAGIVTTFVTLILV--ASLPF	77	
Db	568	GWPNAEL-----TGCCEPIVRVLEWSDIETAIATSCGILVTLFTVFLVIRTPV	622	

```

QY      78 VQDTKKSLLTGTVFFLLGT-LGLFCLVLFACVVVKPDEFSTCASRRFLFGVLFAIFCSCLAA 130
Db      623 VKSSREL-----YIIAGIFLGVVC-PFTLIAKPHTTSCYLQRLLVGSLSSAMCYSAL-- 675
QY      137 HVPALNFLARKNHG-----PR---GWVIETVALLLTLVEVIINTEMLIITLVRGSG 184
Db      676 -VTKTNIARIILAGSKKKIKTRKRPFSMAWAQVIIASILISVOLT-----LVWTLI---- 725
QY      185 EGGPGQNSSAGWAVASP-----CAIANMDFVMALIYVMILLIGLAFLGAW 228
Db      726 -----IMEPPMPILSPSKEVLYLCNTSNLGVAAPGYNGLLIMTSCTYYAF 772
QY      229 ----PALCGRYKRWRKHGFEVLLTTATSVAIHWVMVMTYGKNKOHNSPTWDPTLAT 283
Db      773 KTRNVPAANEAK-----YIAFTMYTTCIIWLAFPVIY-FGSNYKIITTCTFAVSLSVT 824
QY      284 LAANAWAFLVFYVI---PE---VSQYTKSSPEOSYOGD-MYPTRGVGYEYLKEQK-QQS 335
Db      825 VALGCMFTPMMYIIIAIPERNVRSAFTSDVDVRMHVGDKLPGRSNTFLNIFRKKPKGAG 884
QY      336 MEVEN-KAFSEMDEPVAAKRP 354
Db      885 NANSNGKSVSWSEPPGGRQAP 904

RESULT 13
PCT-US91-09423-2
; Sequence 2, Application PC/TUS9109422
; GENERAL INFORMATION:
; APPLICANT: Mulvihill, Eileen R.
; APPLICANT: Hagen, Frederick S.
; APPLICANT: Houamed, Khalel M.
; APPLICANT: Almers, Wolfhard
; TITLE OF INVENTION: G PROTEIN-COUPLED GLUTAMATE RECEPTORS
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/09422
; FILING DATE: 19911212
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,007
; FILING DATE: 18-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/648,481
; FILING DATE: 30-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/626,806
; FILING DATE: 12-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 13952-6PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1199 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```


ATTORNEY/AGENT INFORMATION:
NAME: Blalock, Donna K.
REGISTRATION NUMBER: 38,082
REFERENCE/DOCKET NUMBER: X-8319B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317/277-1090
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1194 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-538-526-1

Query Match 6.38; Score 147; DB 4; Length 1194;
Best Local Similarity 20.68; Pred. No. 2.3e-06;
Matches 104; Conservative 66; Mismatches 192; Indels 142; Gaps 23;
QY 20 GAAQGHVPPCCSOGNPLYYNLCDRSGAWGIVLEAVAGAGIVTFTVLTILV--ASLPPF 77
DB 568 GWW-----PNADLTCCEPIPVRYLEWSNIEPIIAIAFSCGLIVTLFVTLFVLYRDPV 622
QY 78 VQDTKKRSLGTVFFLLGT-LGLFCLVFCVWRPDETCASRRFLFGVLFACFSCIAA 136
DB 623 VKSSSRELC----YIILAGIFLGYVC-PFTLIARPTTTSCYLQRLLVGLSSAMCYSAL-- 675
QY 137 HVFALNFIARKNHG-----PR---GWVFTVALLTLVEIINTFWLIITLVLRSG 184
DB 676 -VTNTRNRIARLAGSKKIKTRKPRFMSAWAQVITIASILISVQLT-----LVVTLI---- 725
QY 185 EGGPGQNSSAGWAVASP-----CAIANMDFVMAIYVMLLLGAFGLAW 228
DB 726 -----TMEPPMPLSYPSKEVYLICNTSLGVAVPLGYNGLLIMSCYYAF 772
QY 229 -----PALCGRYKRWKHGVEVLLTATSAIWMVYTYGNKQHNSTPTWDDPTLAIA 283
DB 773 KTRNVNPFANFEAK-----YIAFTMYTTCIIWLAFVPIY-FGSNYKIITTCFAVLSVT 824
QY 284 LAANAWAFVLYFYI---PE---VSQVTKSSPEQSQGD-MYPTRGVGVETILKEQ----- 331
DB 825 VALGCMETPKMYIIIAKPERNVRSAFTTSDVVRMHVGDGKLPKRSNTFLNIFRRKKAG 884
QY 332 -----KGQSMF-----VENKAFSMDPEVAAKRPVSPSYNGOL 365
DB 885 NANSNGKSVSWSEPGGGOVPKQGHMHRSLVHVKTNETACNQTAIVIKPLTKSYOG-SGKS 943
QY 366 LT-----SVYQPT-----MALMHKVPSEGAYDIILPRATANSOVMG 402
DB 944 LTFSDTSTKTLNVEEEDAQPIRFPSPGSPSMVHRRVPSAATPPPLPHLTAEETPLF 1003
QY 403 SANSTLRAEDMYSAQSHQAATPPK 426
DB 1004 LAEPAL-PKGLPPLPQQOQPPQ 1026

RESULT 9
US-08-041-538-2
Sequence 2, Application US/08041538
Patent No. 5385831
GENERAL INFORMATION:
APPLICANT: Mulvihill, Eileen R
APPLICANT: Hagen, Frederick S
APPLICANT: Houamed, Khaled M
APPLICANT: Almers, Wolfhard
TITLE OF INVENTION: G PROTEIN COUPLED GLUTAMATE
NUMBER OF INVENTIONS: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: CA

COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/041,538
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/648,481
FILING DATE:
APPLICATION NUMBER: US 07/626,806
FILING DATE: 12-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 13952-6-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 206-623-6793
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1199 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-041-538-2

Query Match 6.18; Score 142; DB 1; Length 1199;
Best Local Similarity 22.48; Pred. No. 7.9e-06;
Matches 85; Conservative 54; Mismatches 153; Indels 88; Gaps 19;
QY 20 GAAQGHVPPCCSOGNPLYYNLCDRSGAWGIVLEAVAGAGIVTFTVLTILV--ASLPPF 77
DB 568 GWWPNAEL-----TGCEPIPVRYLEWSNIEPIIAIAFSCGLIVTLFVTLFVLYRDPV 622
QY 78 VQDTKKRSLGTVFFLLGT-LGLFCLVFCVWRPDETCASRRFLFGVLFACFSCIAA 136
DB 623 VKSSSRELC----YIILAGIFLGYVC-PFTLIARPTTTSCYLQRLLVGLSSAMCYSAL-- 675
QY 137 HVFALNFIARKNHG-----PR---GWVFTVALLTLVEIINTFWLIITLVLRSG 184
DB 676 -VTNTRNRIARLAGSKKIKTRKPRFMSAWAQVITIASILISVQLT-----LVVTLI---- 725
QY 185 EGGPGQNSSAGWAVASP-----CAIANMDFVMAIYVMLLLGAFGLAW 228
DB 726 -----TMEPPMPLSYPSKEVYLICNTSLGVAVPLGYNGLLIMSCYYAF 772
QY 229 -----PALCGRYKRWKHGVEVLLTATSAIWMVYTYGNKQHNSTPTWDDPTLAIA 283
DB 773 KTRNVNPFANFEAK-----YIAFTMYTTCIIWLAFVPIY-FGSNYKIITTCFAVLSVT 824
QY 284 LAANAWAFVLYFYI---PE---VSQVTKSSPEQSQGD-MYPTRGVGVETILKEQ-CQS 335
DB 825 VALGCMETPKMYIIIAKPERNVRSAFTTSDVVRMHVGDGKLPKRSNTFLNIFRRKKPG 884
QY 336 MEVEN-KAFSMDPEVAAKRP 354
DB 885 NANSNGKSVSWSEPGGROAP 904

RESULT 10.
US-08-463-642-2
Sequence 2, Application US/08463642
Patent No. 5721107
GENERAL INFORMATION:
APPLICANT: Mulvihill, Eileen R
APPLICANT: Hagen, Frederick S
APPLICANT: Houamed, Khaled M

```

, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: US 08/072,574
, FILING DATE: 04-JUN-1993
, ATTORNEY/AGENT INFORMATION:
, NAME: Reiter, Stephen E.
, REGISTRATION NUMBER: 31,192
, REFERENCE/DOCKET NUMBER: FP41 9772
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 619-546-4737
, TELEFAX: 619-546-9392
, INFORMATION FOR SEQ ID NO: 6:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 879 amino acids
, TYPE: amino acid
, TOPOLOGY: linear
, MOLECULE TYPE: protein
US-08-486-270-6

```

Query Match	6.3%	Score 147;	DB 1;	Length 879;
Best Local Similarity	21.0%;	Pred. No. 1.5e-06;		
Matches	61;	Conservative 55;	Mismatches 106;	Indels 58; Gaps 12;
Qy	48	AWGIVLEAVACAGIVTTFVLTIIIVA--SLPFDVDTKRSLLGTQVFLLTGTLGL-FCLV 104		
Db	575	AWAIGPVYTIACLGMCTCMVVTVFIRKNNTPLVRASGRE-----LCYTILFGVGLSYCMT 629		
Qy	105	FACVVKPDFSTCASRRFLFGVLFAICHSCLAHHVFNLFAR----KNHGPRGWVI--- 156		
Db	630	FFFTAKSPVICALURRLGLSSFAICYALSALTKT--NCIARIEDGVKGNAQRPKFI SPS 686		
Qy	157	--FVALLLTVEVINTEWLIITLVRGSGGGQGQNSSAGAWAVS-----PCAATNM 207		
Db	687	SQVFCIGLILVQIVMVSWMIL-----EAPGTRRYTIAEKRET VILCKNVKDS 735		
Qy	208	DFVMALIVMLLLLGAFELGAMPALCGRYK-RWRK-----HGVFVLLTTATSAIVWVMI 260		
Db	736	SMLSISLYDVILVI-----LCTVYVPKTRCPENFEAKFIGFTMYTTCIWLAF 786		
Qy	261	VMYTYGNKHNSPTWDDPTLAIALAANA WAFVLPYVIP EV 300		
Db	787	P IFVYTSDDRVQVQ-----TTMCISIVSLSGFVVLGCIFAPKV 823		

RESULT 7
US-08-367-264-6
; Sequence 6, Application US/08367264
; Patent No. 6001581
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrle
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen
; APPLICANT: Pontsler, Aaron
; APPLICANT: Johnson, Edwin C.
; APPLICANT: Hess, Stephen D.
; TITLE OF INVENTION: HUMAN METABOTROPIC GLUTAMATE RECEPTORS,
; NUCLEIC ACIDS ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/367,264
; FILING DATE: 02-JUN-1994

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/072,574
; FILING DATE: 04-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: FP41 9772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-4737
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-367-264-6

Query Match 6.3%; Score 147; DB 3; Length 879;
Best Local Similarity 21.8%; Pred. No. 1.5e-06;
Matches 61; Conservative 55; Mismatches 106; Indels 58; Gaps 12;

Qy 48 AWGIVLEAVAGAGIVTFVLTIIIVA--SLPFDQDTKRSLLGTQVFLTLGLTGL--FCLV 104
Db 575 AWAICPVTIACLGPMCTCMVVTVFIKHNTPLVKASGRE----LCYTLLFGVGLSYCMT 629

Qy 105 FACVVKPDFSTCAARRFLGVFLFAICFSCLAAHVFAINFLAR----KNHGPRGWI--- 156
Db 630 FFTAKPSPTICALRRGLGSSFAICYALLTKT---NCIARIFDGVKNGAQRPKFISPS 686

Qy 157 --FTVAALLTLVEIVINTEFWLIITLVRSGGGPGQSSAGWAVAS-----PCAIANM 207
Db 687 SQVFICLGILLVQIVMVSVWLIL-----EAPGTRRYTLAEKRETVILKCNVKDS 735

Qy 208 DFVMALIVMULLGAFGLGAWPALCGRYK--RWK-----HGVFVLLTTATSVAIWVWI 260
Db 736 SMLISLYDVILVT-----LCTVYAFKTRKCPENFNEAKFIFTMTYTCIIWLAF 786

Qy 261 VMYYIGNKQHNPSPTWDDPTLAIALAANAFAVLFVIVEV 300
Db 787 PIFVYTTSSDYRVQT---TTMCISVLSLGVFVVLGCLFAPKV 823

RESULT 8
US-08-538-526-1
; Sequence 1, Application US/08538526
; Patent No. 6303751
; GENERAL INFORMATION:
; APPLICANT: Burnett, J. Paul
; APPLICANT: Mayne, Nancy G.
; APPLICANT: Sharp, Robert L.
; APPLICANT: Snyder, Yvonne M.
; TITLE OF INVENTION: Human Metabotropic Glutamate Receptor
; TITLE OF INVENTION: and Related DNA Compounds
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Division/DKB
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/538,526
; FILING DATE: October 3, 1995
; CLASSIFICATION: 530

```

; APPLICATION NUMBER: 08/337,797
; FILING DATE: NO. 6103475ember 14, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaylo, Paul J.
; REGISTRATION NUMBER: 36,808
; REFERENCE/DOCKET NUMBER: X-9431
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 276-0756
; TELEFAX: (317) 276-3861
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 872 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-258-523-2

Query Match 6.6%; Score 153; DB 3; Length 872;
Best Local Similarity 22.2%; Pred. No. 3.3e-07;
Matches 75; Conservative 50; Mismatches 143; Indels 70; Gaps 14;

Qy 6 ALVNCGLP--LFUFGAWAGHVPPGCSQGLNPLYNLCORSAGWIVLEAVAGAGIVT 63
Db 549 SLTGCFFLPQEIYRWGDWAVGPVTIAC-----LGA-LAT 582

Qy 64 TFVLTITLVASLPVQDTKKRSLTGTVFFLLGLTGL-FCLVFACVVKVPDEFSTCASRRFL 122
Db 583 LFVLGVFVRINATPVVWASGREL-----CYLLGGVFLCYCMTFIFIAKPSGTGVCALRRLG 638

Qy 123 FGVLFACFCFCLAAHVFALFNLRKHGPRGW-----IFTVALLLTIV--EVIINT 172
Db 639 VGTAFSCVYSALLTKT---NRIARIFGAREGAORPRFISPASOVAICLALISQQLIVV 695

Qy 173 EWLITLVRSGEPPGCGNSAGWAVASPAICANMDFMALIYVMLLLGLGAFGAWPALC 232
Db 696 AWWVEAPGKGKETAPERRE-----VTLRCNHRDASMLGSLAYNVLLI-----ALC 742

Qy 233 GRYK-RWRK-----HGVFVLLTATSVAIWVWVIMVYGNKGNKHNPTWDDPTLATALA 285
Db 743 TLVAFKTRKCPENFNEAKFGFTMYTTCIIWLAFLPFIYVTSDDYRVQT---TTMCVSVS 799

Qy 286 ANAWAFVLFVIVPEVSQVTKSSPEQSQGDMYPTRGVG 323
Db 800 LSG-SVVLGCLFAPKHLIIILFQPKNVVSHRAPTSREG 836

RESULT 5
US-08-072-574-6
; Sequence 6, Application US/08072574
; Patent No. 5521297
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrle
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen
; APPLICANT: Pontsler, Aaron
; TITLE OF INVENTION: HUMAN METABOTROPIC GLUTAMATE RECEPTORS,
; NUCLEIC ACIDS ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 00719
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/072,574

; FILING DATE: 19930604
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter, Stephen E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9383
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 213-622-7700
; TELEFAX: 213-489-4210
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 879 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-072-574-6

Query Match 6.3%; Score 147; DB 1; Length 879;
Best Local Similarity 21.8%; Pred. No. 1.5e-06;
Matches 61; Conservative 55; Mismatches 106; Indels 58; Gaps 12;

Qy 48 AWGIVLEAVAGAGIVTTFVLTIIIVA--SLPFVQDTKKRSLGTVQVFFLLGLTGL-FCLV 104
Db 575 AWAIGPVTIACLGPMCTCMVVTVEIKHNNTPLVKASGRE-----LCYILLFGVGLSYCMT 629

Qy 105 FACVVKDFSTCASRRFLPGVLFPAICFSCSLAAHVFAFNFLAR-----KNHGPRGWV---- 156
Db 630 FFFIAKPSVICALURRLGLSGSSFAICYASALLTKT---NCTARIFDGKVGKGAQRKPFISPS 686

Qy 157 -FTVALLLTIV--EVIINTWLTITLVRSGBGGPGQNSAGWAVAS-----PCAIANM 207
Db 687 SQVFICGLILVQIVWVSVMLIL-----EACGTRRYTLAEKRETVILKCNKVD 735

Qy 208 DFVMAIYVMLLLGLGAFGAWPALCGRYK-RWRK-----HGVFVLLTATSVAIWVWI 260
Db 736 SMLISLYDVILVI-----LCTVYAFKTRKCPENFNEAKFGFTMYTTCIIWLAF 786

Qy 261 VMTYGNKHNKHNPTWDDPTTALIAAANAWAFVLFYVPEV 300
Db 787 PFIYVTSDDYRVQT---TTMCISVSLSGFVVVLGCLFAPKV 823

RESULT 6
US-08-486-270-6
; Sequence 6, Application US/08486270
; Patent No. 5807689
; GENERAL INFORMATION:
; APPLICANT: Daggett, Lorrle
; APPLICANT: Ellis, Steven B.
; APPLICANT: Liaw, Chen
; APPLICANT: Pontsler, Aaron
; APPLICANT: Johnson, Edwin C.
; APPLICANT: Hess, Stephen D.
; TITLE OF INVENTION: HUMAN METABOTROPIC GLUTAMATE RECEPTORS,
; NUCLEIC ACIDS ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,270
; FILING DATE: 02-JUN-1994
; CLASSIFICATION: 435

Db 271 -----DAECKPQLVKKSYGVENRAYSQEE-----ITQGFETGDTLYAPYS-THFQ 315
QY 377 LMHKVPSEGAYDILPRATA 396
Db 316 LQNOPPOK-----EFSIPRAHA 332

RESULT 2

US-09-188-930-123
Sequence 123, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 123
LENGTH: 68
TYPE: PRT
ORGANISM: Human
US-09-188-930-123

Query Match 7.3%; Score 169; DB 4; Length 68;
Best Local Similarity 54.3%; Pred. No. 1.3e-10;
Matches 38; Conservative 11; Mismatches 17; Indels 4; Gaps 2;

QY 66 VLTI-ILVASLPFVQDTKKSLGTQVFLGLTGLFCLVACVVKPDEFSTCASRRFLG 124
Db 1 MLTPLVLVCK--VQDSNRKRMKLPOTFLGLVGLFGLTFAFTIGLDGSGTGPTREFLG 57
QY 125 VLFAICFSCSL 134
Db 58 ILFSICFSCSL 67

RESULT 3

US-08-337-797A-2
Sequence 2, Application US/08337797A
Patent No. 6017697
GENERAL INFORMATION:
APPLICANT: Burnett, J. P.
APPLICANT: Mayne, Nancy G.
APPLICANT: Sharp, Robert L.
APPLICANT: Snyder, Yvonne M.
TITLE OF INVENTION: EXCITATORY AMINO ACID RECEPTOR PROTEIN
TITLE OF INVENTION: AND RELATED NUCLEIC ACID COMPOUNDS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: United States of America
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/337,797A
FILING DATE: No. 6017697ember 14, 1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:

NAME: Gaylo, Paul J.
REGISTRATION NUMBER: 36,808
REFERENCE/DOCKET NUMBER: X-9431
TELECOMMUNICATION INFORMATION:
TELEPHONE: (317) 276-0756
TELEFAX: (317) 276-3861
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 872 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-337-797A-2

Query Match 6.6%; Score 153; DB 3; Length 872;

Best Local Similarity 22.2%; Pred. No. 3.3e-07;
Matches 75; Conservative 50; Mismatches 143; Indels 70; Gaps 14;
QY 6 ALVMCLGLP--LFLFPGAWAQGHVPPGCSQGLNPLYYNLCDRSGAWGIVLEAVAGAGIVT 63
Db 549 SLTGCFELPQEIYRWGDANAVGPVTIAC-----LGA-LAT 582
QY 64 TFVLTIILVASLPFVQDTKKRSLGTQVFLGLTGL-LFCLVACVVKPDEFSTCASRRFL 122
Db 583 LFLVGLVFVRHNATPVVKASGREL---CYLLGGVFLCYCMTEFIATKPSGTGVCALRRLG 638
QY 123 FGVLFATCFSCSLAAHVFAFNFLARKNHGPRGW-----IFTVALLTLV--EVIINT 172
Db 639 VGTAFSVYCSALLTKT---NRIARIFGGAREGAQRPFISPASQVAICLALISGQLLIV 695
QY 173 EWLITILVRSGEGGQGNSSAGWAVASPCAIAIMDFVMAIYVMLLLGLGAFLGAWPALC 232
Db 696 AMLVVEAPGPGKETAPERRE---VVTLCRHRDASMLGSLAYNVLLI-----ALC 742
QY 233 GRVK-RWRK-----HGVFVLLTATSAIVLVWVWVWYTYGNKQHNSTPDDTLAIALA 285
Db 743 TLVAFTRKCPENFENAKFTGFTMYTTCIIWLAFLPIFYVTSSDYRVQT---TTMCVSVS 799
QY 286 ANAWAFVLFVYIPEVSQVTKSSPEQSYQGDMYPTRGV 323
Db 800 LSG-SVVVLGCLFAPKLHIILFQKNNVSHRAPTSRFG 836

RESULT 4

US-09-258-523-2
Sequence 2, Application US/09258523
Patent No. 6103475
GENERAL INFORMATION:
APPLICANT: Burnett, J. P.
APPLICANT: Mayne, Nancy G.
APPLICANT: Sharp, Robert L.
APPLICANT: Snyder, Yvonne M.
TITLE OF INVENTION: EXCITATORY AMINO ACID RECEPTOR PROTEIN
TITLE OF INVENTION: AND RELATED NUCLEIC ACID COMPOUNDS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: United States of America
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/258,523
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

Search completed: September 16, 2002, 09:32:47
Job time: 206 sec

RESULT 15

thio-disulfide interchange protein, probable [imported] - Caulobacter crescentus

Query Match

12 GLPLFLFPG-AWAOGHVPPGCSOGI.NPLYYNI.CDRSGAWGTVI.---EAVAGA----- 59

Best Local Similarity

Matches 105; Conservative 57; Mismatches

12 GLPIELFPG-AWAQGHVPPGCSQGLNPIYYNLCDRSGAWGTVI.---FAVAGA----- 59

[illegible]

223 GLTLTKPGYDFVGGSTPPAELVG-----LATKAGAEVTATAG_EALAGASGLGPPA 276

60 -----GIVTTEVIT III VASI BEVADTKKRSI I CTAVEEI I CTICI ECI VEACIAWKRRDS 114

[illegible]

277 EAAPGALA-----GGLAGALLFAFLG--GLILNLMPCVF-PVLS 313

115 TCASB-----DET ECVI EAT TCRCT A AUVWAY NET A B YNNUCDBCW-----154

Y	ICASK	REFUGEE	SC	AAHV	FA	NE	EA	KN	NG	RGW	---	134
	:			:	:	:	:	:	:	:		

314 MKAASLAGHAHEAPKARLQGLAFLGVV--ATFLALAGALLAV---RAGGAAVGWGFQL 367

165 - - - - - VY CMVRY Y CMY WCVYT YANMELBY Y YMY VND
C C

133 ---VIFVALLTILVEVIINIEWLIIILVRG----SGEGGPGQ---NSSAGWAVASPC 202

368 QSPLVIAGLALLMLVALNMSGIFEIGTSVQGVGAGASAKGVSGAFFTGALAVVVAAPC 427

U	T	C
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

203 ATANMDFVM--ALTYVMLLLGAF LG-----AWPALCGRYKRWRKHGVFV----- 245

428 TAPFMAGALGYALTQPPVIALGVFLALALGFAAPEFVAVTFIPGALKLLPRPGDWMEVLKK 487

246 -LLTATSAIWWIV-----MYTYGNKQHNSPTWDDPTLAI 282

488 GLAFPMYGAALWLWVFAQQAGPIALGQLLVAGVLAAFGAWLYCLAQARRAVKGS--AV 545

283 ALAANAWAFVLFYVPEVSQVTRSSPEQSYQGDMYPTRGCVGYETILKEQGQSMFVENKA 342

546 SMILGLLAVVGALALAAASALSAKPPVAAAEASTPSGPGLTAEA-WSPEKVOALOEGRP 604

343 FSMD-----EPVAAKRP--VSPYSGYNGQLLTSVYQPTMALMHKV----- 381

A:Reference number: I49142; MUID:95239344
A:Accession: I49142
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-908 <RES>
A:Cross-references: EMBL:U17252; NID:g854728; PIDN:AAA68149.1; PID:g8547292
C:Genetics:
A:Gene: mglu8
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: neurotransmitter receptor

Query Match 5.4%; Score 125; DB 2; Length 908;
Best Local Similarity 20.4%; Pred. No. 0.03;
Matches 84; Conservative 75; Mismatches 144; Indels 108; Gaps

QY 19 PGAWAOGHVPGGCSQGLNPLLYNLC-----DRSGA-----WGIVLEAVA 57
DB 533 PCWHGRC-EGYNYQVDELSCELCPLDQRPNNRTGCRIPILKLEWISPAWVPLIA 591
QY 58 GAGIV-TTFVL-TIILVASLPFFQDTKKRSLGTQVFFLLGTGLF--CLVFACVVKPD 112
DB 592 ILGIIATTEVIVTFVRYNDPIVRAS-----GRELSYLLT-GIFLCYSITFLMIAAD 644
QY 113 FSTCASRRFLGVLFALCFSCLAHVFALNFLAR-----KNHGRGVWIFTVA 160
DB 645 TIICSFRRIFGL--GMCF5-YAALLTKNRIHRIEFOGKKSVPATKFI5PASQLVIT5 701
QY 161 LL-LTLVEVIINTEWLI-----TLVRGSGGGGQGNSSAGWAVASPCATANMDFVMALIY 215
DB 702 LISVQLLGVFV---WFVDDPHHTIIDYGEORTLDPENARG---VLKCDISDLSLCSLGY 755
QY 216 VMLLLGLAFLGAMPALCGRYKRWKHGV-----FVLLTATSAIVWVWVIMVT 264
DB 756 SILLW-----TCTVY-AIKRGVPTFNEAKPIGFTMYTTC--IILAFIPIF- 801
QY 265 YGNKQHSPTW-DPDLTALAIANAFAVLFFYVPEVSQVTKSSPQ5YOGOMYPTRGVG 323
DB 802 FGTAQSAEKWYIQTTTLT5VMSLSASVSLGMLYMPKV-VYIIFHPQBNVQ-----KKRS 855
QY 324 YETILKEQKQSGMFVENKAF5MDEPVAAKRPVSPYSGYNCQLLT5VYQPT5 374
DB 856 FKAVVTAATWQSLIQ-----GNDRP-----NGEVKSELCE5LE 890

RESULT 13
I59362
calcium/polyvalent cation-sensing receptor precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C:Accession: I59362; A55594
R:Ruat, M.; Molliver, M.E.; Snowman, A.M.; Snyder, S.H.
Proc. Natl. Acad. Sci. U.S.A. 92, 3161-3165, 1995
A:Title: Calcium sensing receptor: molecular cloning in rat and localization
A:Reference number: I59362; MUID:95241465
A:Accession: I59362
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1079 <RES>
A:Cross-references: EMBL:U20289; NID:g790578; PIDN:AAC52195.1; PID:g790579
A:Experimental source: striatal
R:Riccardi, D.; Park, J.; Lee, W.S.; Gamba, G.; Brown, E.M.; Hebert, S.C.
Proc. Natl. Acad. Sci. U.S.A. 92, 131-135, 1995
A:Title: Cloning and functional expression of a rat kidney extracellular
A:Reference number: A55594; MUID:95116508
A:Accession: A55594
A:Molecule type: mRNA
A:Residues: 1-133,'X',135-1079 <RTC>
A:Experimental source: kidney
A:Cross-references: GB:U10354
A:Experimental source: kidney
C:Keywords: calcium; glycoprotein; phosphoprotein; transmembrane protein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:187-212/Region: hydrophobic

[illegible]

A:Title: Calcium sensing receptor: molecular cloning in rat and localization to nerve terminal
A:Reference number: I59362; MUID:95241465
A:Accession: I59362
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1079 <RES>
A:Cross-references: EMBL:U20289; NID:g790578; PIDN:AAC52195.1; PID:g790579
A:Experimental source: Striatum
R:Riccardi, D.; Park, J.; Lee, W.S.; Gamba, G.; Brown, E.M.; Hebert, S.C.
Proc. Natl. Acad. Sci. U.S.A. 92, 131-135, 1995
A:Title: Cloning and functional expression of a rat kidney extracellular calcium/polyvalent cation receptor
A:Reference number: A55594; MUID:95116508
A:Accession: A55594
A:Molecule type: mRNA
A:Residues: 1-133, 'X', 135-1079 <RIC>
A:Cross-references: GB:U10354
A:Experimental source: Kidney
C:Keywords: calcium; glycoprotein; transmembrane protein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:187-212/Region: hydrophobic

```

Qy   78  VQTKKRSLGTVFFLLGT-LGLFCLVACVVVKPDEFSCASRRFLFGVLFAICFSCLAA 136
      |: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    623 VKSSSREL-----YIILAGIFLGVC-PETLIAKPTTTCYLQRLLVLGSAMCYAL-- 675

Qy   137 HVFALNPLARKNHG-----PR---GWVFIVALLTLVEVINTEWLIITLVRSG 184
      |: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    676 -VTKTNIARILAGSKKKICTRKPREMSAWAQVIIASILISVOLT-----LVLTILI---- 725

Qy   185 EGGPOGNSSAGWAVASP-----CATANDFMALIYVMLLCGAFLGAW 228
      |: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    726 -----IMPEPPILSYPSIKEVYLICNTSNLGVAPGYNGLLIMSTYYAF 772

Qy   229 -----PALCGRYKRWRKHGVFLLTTATSVIAIWMVIMTYTGKNHNSPTWDPTLAIA 283
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    773 KTRNVPANFNPAKEAK-----YIAFTMYTCTIIWLAFPVIY-FGSNKIITTCFVSLSVT 824

Qy   284 LAANAFADVLFYVI-----VSQVTKSQSQGSD-MYPTRGVGVETILKEOK-QQS 335
      |: : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    825 VALCGMFTPKNYIIIAKPENRVSAFTTSDVVRMHVGDKLPKCRNSTELNIPFRKKPGAG 884

Qy   336 MFVEN-KAFSMDEPVAAKRP 354
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    885 NANSNGKSVSWSEPGGRQAP 904


RESULT 10
$?1376
glutamate receptor homolog - cherry salmon
C:Species: Oncorhynchus masou (cherry salmon)
C:date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 01-May-1998
C:Accession: $?1376
R:Kubokawa, K.; Miyashita, T.; Nagasawa, H.; Kubo, Y.
```

Qy	253	VAIWVWIVMYVTYGNKOHNSPTWDDPTLATAAANAWAFVLYVI---PE-----VSQVTKS	306
		: : : :	: : : :
Db	805	CIIWLAEVPYI-FGSNKKIITTSFSSVSLSTVALGCMFSFKYIILAKPVRNVRSFAFTTS	863
		: : : :	: : : :
Qy	307	SPEQSIOGD-MYPTRGVGYETILKEOGSGSMFVEN---KAFSDPEVAAKRPVSPYSQYN	362
		: : : :	: : : :
Db	864	DVVRMHVYDGGNVACRSSLMDMPKRKNKS-----ENSGKSVSWSEPGA--RHHPAGGEHN	917
		: : : :	: : : :
Qy	363	GQL-----LTSVTOPT--EMAL-----MHKVPSSEGAYDIH	390
		: : : :	: : : :
Db	918	WORLSVVHKRQEACCSNQMAVIRPLINTYHTNGTSSNAMEFSDLSTKTLYNVAEEDESDLV	977
		: : : :	: : : :
Qy	391	-----LPRATANSQVMGNSANSTLRAD	412
		: : : :	: : : :
Db	978	RYNPLSPHHMAHGQINATTGGPMKEVD	1005
		: : : :	: : : :
RESULT	11		
T30806		metabotropic glutamate receptor 7 - Fugu rubripes (fragment)	
C:Species:	Fugu rubripes		
C>Date:	02-Sep-2000	#sequence_revision	02-Sep-2000 #text_change
C:Accession:	T30806		
R:Cottage,	A.J.; Clark, M.; Hawker, K.; Umrانيا, Y.; Wheller, D.; Bishop, M.		
FEBS Lett.	443, 370-374, 1999		
A>Title:	Three receptor genes for plasminogen related growth factors in the		
A:Reference number:	Z20880; MUID:99148833		
A:Accession:	T30806		
A>Status:	preliminary;		
A:Molecule type:	DNA		
A:Residues:	1-551 <COT>		
A:Cross-references:	EMBL:AJ010317; NID:e1355235; PID:e1355236; PIDN:CAA090808		
C:Genetics:			
A:Gene:	GRM7		
A>Note:	intron positions not resolved (incomplete sequence)		
C:Superfamily:	metabotropic glutamate receptor 4		

A: residue type: MKRK
 A: residues: 1-1218 <KUB>
 C: keywords: glycoprotein; phosphoprotein
 F: 603-625/Domain: transmembrane #status predicted <TM1>
 F: 640-660/Domain: transmembrane #status predicted <TM2>
 F: 672-690/Domain: transmembrane #status predicted <TM3>
 F: 717-737/Domain: transmembrane #status predicted <TM4>
 F: 761-782/Domain: transmembrane #status predicted <TM5>
 F: 796-817/Domain: transmembrane #status predicted <TM6>
 F: 826-850/Domain: transmembrane #status predicted <TM7>
 F: 104, 233, 403, 525, 757/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F: 636, 699, 961/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
 F: 705/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F: 892/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predicted

Query Match	5.7%;	Score 132.5;	DB 2;	Length 551;
Best Local Similarity	19.3%;	Pred. No. 0.0044;		
Matches 57;	Conservative 64;	Mismatches 123;	Indels 51;	Gaps
Qy	46	SGANGIVLEAVAGAGIVTT--FVLITILVASLPFVODTKKRSLLGTQVFFLL--CTLGLF	l01	
	: : : :	: : : :	: : : :	:
Db	242	SSPAIIPVLAIIIGLITATGVIATIRFNDDTPIVRAS-----GRELSYVLLTGIFLIY	295	
Qy	102	CLVFACVVKPDEFSCASRRRLFGVLAFCFSCLAA---HVFALNLFARKNHGPRGWIFT	158	
	: : : :	: : : :	: : : :	:
Db	296	LITELMAEPSVAACAERLLFLGLMCISYSAMLTNTRIYRIFEGOKSKVTVPKFISPT	355	
Qy	159	VALLTIIVEIINTEWLII-----TLVRSGEGGPQGNSAGAWASPACAIANMDEV	210	
	: : : :	: : : :	: : : :	:
Db	356	SOLIITFILTSVQLLGVIWFEWGVMPPHTIIDEEQQPPNPFAERG---VLKCDMSDLSLI	412	
Qy	211	MALIYVMLLLIGLAFGAMPALCGRYKRWRKHGV-----FVLLTTATSVIAVWMV	259	
	: : : :	: : : :	: : : :	:
Db	413	LCLSLSYLMI-----TCIVY-AIKSRGVPETFNEAKPIGETMTWTC---IVWLAF	459	
Qy	260	IVWTYGNKOHNSPTW--DDPTLAIALAANAAPLVYEIVTPVESQVTKSPESQSYQ	313	
	: : : :	: : : :	: : : :	:
Db	460	VPIC-FCTAOSTEKMFQTQTTTLTVSMSLSATVSLGLMYLPKV-YVITIHFPEQNQV	512	

RESULT 12
149142
metabotropic glutamate receptor 8 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1996
C:Accession: 149142
R:Duvoisin, R.M.; Zhang, C.; Ramonell, K.
J. Neurosci. 15, 3075-3083, 1995
A: Title: A novel metabotropic glutamate receptor expressed in the retina and

A;Cross-references: GB:L08132; NID:g290215; PIDN:AAA72332.1; PID:g290216
C;Genetics:
A;Gene: FlyBase:Dvir/boss
A;Cross-references: FlyBase:FBgn0013105
A;Introns: 14/3; 175/3; 324/3; 808/3
C;Keywords: transmembrane protein

Query Match 6.3%; Score 146; DB 2; Length 893;
Best Local Similarity 20.8%; Pred. No. 0.0006;
Matches 71; Conservative 58; Mismatches 136; Indels 76; Gaps 15;

QY 34 GLNPYYNLDRSCA--WGIVLEAVAGAGIVTFVLTIIILVASLPFFVDTKKRSLLG--- 88
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 503 GENRYPFLFDGESVMFWRIKLDTWATTAALIGLIATLAILVFI---VVRISLGDNVF 559
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 89 -----TQVFFLLGTGLGFC--LVFAC-----VVKPDFST-----CASRRFLFGVLV 127
| | | | | : | : | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 560 EGNPVTSEILLILLSILVFCFSVPESMEYVGQRNSHVTFFEDVHTLNTLCGRVRFIMTLVY 619
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 128 AICFSCIAAHVFAL-----NFLARNKHGPRGWIFTVALLTLVEIINTEWLITLV 180
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 620 CFVPSLLLCRAVMLASIGEGFLSHVN---GYQAIICVLSVFVGMSVQLLVV--M 673
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 181 RSGEGPGQNSSAG-WAVASPCAANDFPMALIYVMIILLGAFLGANWALCALCRYKRW 239
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 674 HLASESVSCENIYYGRW-----LWGLLAYDFLLCLSLVSLVPFYQRSORNY 720
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 240 KHGEVLLTTATSAVAIWVVMTYTYKNQHNSPTWDDPTLATLAANAFAVLFYVIP- 298
| | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 721 E-GILVIGAVLIIILWSWIALSFMFDE-----WRDAIPICMGQASGNAVLVGILIPR 773
| | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 299 -----EVSQVTKSSPEQS---YQDMYPTRGVGYETI 327
| | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 774 TFLIVRGERSDIAQAALPSLTSLAF AQNNQYSSESQSVYECV 814
| | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 9
A41939
G protein-coupled glutamate receptor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C;Accession: A41939; S15362
R;Houamed, K.M.; Kuljper, J.L.; Gilbert, T.L.; Haldeman, B.A.; O'Hara, P.J.; Mulvihill,
Science 252, 1318-1321, 1991
A;Title: Cloning, expression, and gene structure of a G protein-coupled glutamate rec
A;Reference number: A41939; MUID:92022526
A;Accession: A41939
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-1199 <HOU>
A;Cross-references: CB:M61099; NID:g397806; PIDN:AAA19497.1; PID:g204460
A;Experimental source: Cerebellum
A;Note: sequence extracted from NCBI backbone (NCBIP:60785)
R;Masu, M.; Tanabe, Y.; Tsuchida, K.; Shigemoto, R.; Nakanishi, S.
Nature 349, 760-765, 1991
A;Title: Sequence and expression of a metabotropic glutamate receptor.
A;Reference number: S15362; MUID:91156047
A;Accession: S15362
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1199 <MAS>
A;Cross-references: EMBL:X57569; NID:g56646; PIDN:CAA40799.1; PID:g56647
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 6.1%; Score 142; DB 2; Length 1199;
Best Local Similarity 22.4%; Pred. No. 0.0018;
Matches 85; Conservative 54; Mismatches 153; Indels 88; Gaps 19;

QY 20 GAWAQGHVPPCGSGLNPPLYNLCDRSGANGWIVLEAVAGAGIVTFVULTIYL--ASLPF 77
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 568 GWPNPAEL-----TGCEPIPVRYLEWSDIESTIAIAFSLGILVTLFTLVLYRTFPV 622
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Best Local Similarity 19.8%; Pred. No. 5.1e-05;
Matches 69; Conservative 63; Mismatches 130; Indels 86; Gaps 14;

QY 32 SQGLNPLYNLCDRSGA--WGIVLEAVAGAGIVTTFTVLITILVASLPFVQDTKKRSLLG- 88
Db 504 SAGDNNRRYPFLDGESVMFWRKMDTWATGLTAAILGLIATLAILVFI---VVRISLGD 560

QY 89 -----TQVFLLGLTGLFCLVF-----ACVVKPDFST-----CASRRFLFGV 125
Db 561 VFEGNPTTSILLLLSLLVFCSEVPYSIEYVGEQRNSHTVFEDAQTLNTLCVAVRFIMTL 620

QY 126 LFAIFCSCLAAHFAL-----NFLARKNHGPRGWVFTVALLLLTILVEIINTEWLIIT 178
Db 621 VYCFVFSLLLCRAVMLASIGSEGGFLSHVN---GIYQAVICAFSVVAQVGMVSQVLLVVM 676

QY 179 LVKSGEGGPGQNSSAGWAVASPCATANMDF-----VMALIYVMLLLLLGAFIGAWPALC 232
Db 677 HV-----ASETVCENIYYGRWLWGLLAYDFALLCCVCGALI---PSIY 716

QY 233 GRYKRRKHGCVFVLLTTATSVAIWVWVIMVYTYGNKQHNSTPTWDDPTLATAAANAWAFV 292
Db 717 RSNQRTRE-GILIVIGSVLIMVIAWIALSLFGE-----WRDAIPIGLQASGMAVL 769

QY 293 LFYVIP-----EVSQVTKSPSQS-----YQGDWYPTRGVYETI 327
Db 770 VGILIPRTFLIVRGIERSDIAQALPSLTSIAFAQNQYSSESQSVYECV 817

RESULT 5
S26740
gene boss protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 16-Feb-1997
C:Accession: S26740
R:Kramer, H.
submitted to the EMBL Data Library, September 1990
A:Reference number: S26740
A:Accession: S26740
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-896 <KRA>
A:Cross-references: EMBL:X55887
C:Genetics:
A:Gene: FlyBase:boss
A:Cross-references: FlyBase:FBgn0000206
A:Introns: 18/1; 175/1; 325/1; 811/3

Query Match 6.7%; Score 157; DB 2; Length 896;
Best Local Similarity 20.5%; Pred. No. 7.4e-05;
Matches 87; Conservative 65; Mismatches 166; Indels 106; Gaps 19;

QY 32 SQGLNPLYNLCDRSGA--WGIVLEAVAGAGIVTTFTVLITILVASLPFVQDTKKRSLLG- 88
Db 504 SAGDNNRRYPFLDGESVMFWRKMDTWATGLTAAILGLIATLAILVFI---VVRISLGD 560

QY 89 -----TQVFLLGLTGLFCLVF-----ACVVKPDFST-----CASRRFLFGV 125
Db 561 VFEGNPTTSILLLLSLLVFCSEVPYSIEYVGEQRNSHTVFEDAQTLNTLCVAVRFIMTL 620

QY 126 LFAIFCSCLAAHFAL-----NFLARKNHGPRGWVFTVALLLLTILVEIINTEWLIIT 178
Db 621 VYCFVFSLLLCRAVMLASIGSEGGFLSHVN---GIYQAVICAFSVVAQVGMVSQVLLVVM 676

QY 179 LVKSGEGGPGQNSSAGWAVASPCATANMDF-----VMALIYVMLLLLLGAFIGAWPALC 232
Db 677 HV-----ASETVCENIYYGRWLWGLLAYDFALLCCVCGALI---PSIY 716

QY 233 GRYKRRKHGCVFVLLTTATSVAIWVWVIMVYTYGNKQHNSTPTWDDPTLATAAANAWAFV 292
Db 717 RSNQRTRE-GILIVIGSVLIMVIAWIALSLFGE-----WRDAIPIGLQASGMAVL 769

QY 293 LFYVIP-----EVSQVTKS-----SPEQSYQGDWYPTRGV---YETILKEQG 333

Db 770 VGILIPRTFLIVRGIERSDIAQALPSLTSIAFAPEQSI-----LLGTGSYVECVNPMARR 823
QY 334 QSMFVENKAFSDEPVAAKRPVSPYSGYNQQLTTSVQPTFEMALMHKV-----PSEGAIDI 389
Db 824 CSODEVNHQSPSEIPTPLRGGGP---RRQOFFANLRQMP-TLIHNVHLDRSRVPPDP 879

QY 390 ILPR 393
Db 880 RCPR 883

RESULT 6
JH0561
metabotropic glutamate receptor 2 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 13-Sep-1998
C:Accession: JH0561
R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
Neuron 8, 169-179, 1992
A:Title: A family of metabotropic glutamate receptors.
A:Reference number: JH0561; MUID:92110002
A:Accession: JH0561
A:Molecule type: mRNA
A:Residues: 1-872 <TAN>
A:Experimental source: brain
C:Comment: This protein is coupled to a G protein and evokes a variety of functions b
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane p
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-872/Product: metabotropic glutamate receptor 2 #status predicted <MET>
F:568-590/Domain: transmembrane #status predicted <TRI>
F:605-625/Domain: transmembrane #status predicted <TII>
F:637-655/Domain: transmembrane #status predicted <III>
F:680-700/Domain: transmembrane #status predicted <TIV>
F:726-747/Domain: transmembrane #status predicted <TRV>
F:761-782/Domain: transmembrane #status predicted <TVI>
F:795-819/Domain: transmembrane #status predicted <VII>
F:203,286,338,402,547/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:601,675,827,837,843/Binding site: phosphate (Ser) (covalent) #status predicted
F:832/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 6.5%; Score 152; DB 2; Length 872;
Best Local Similarity 21.9%; Pred. No. 0.00019;
Matches 74; Conservative 51; Mismatches 143; Indels 70; Gaps 14;

QY 6 ALVMCLGLP--LFLFPGAWAQGHVPPGCSQGLNPLNYNLCDRSGAMGIVLEAVAGAGIVT 63
Db 549 SLTGCELPQEIYRWGDAMAVGPVTIAC-----LGA-LAT 582

QY 64 TFLVTILVASLPFVQDTKKRSLLGTVQVFLGLTGL-FCLVFACVVKPFDFTCASRRFL 122
Db 583 LFLVGFVRHNATPVVKASGREL---CYILGGVFLCYCMTEVFIKAPSTAVCTLRLRG 638

QY 123 FGVLEAIFCSCLAAHFALNFARKNHGPRGWV-----IFTVALLTLV---EVLINT 172
Db 639 LGTAFSVCSYALLTKT---NRIARIFGAREGAGRPRFTSPASQVAICLALISQQLLIVA 695

QY 173 EWLIITLVKSGEGGPGQNSSAGWAVASPCAIANMDFVMALIYVMLLLLLGAFIGAWPALC 232
Db 696 AWLVEAPGTGKETAPERRE---VVTLCNRHDSMLGSLAYNLLI-----ALC 742

QY 233 GRYK-RWRK-----HGVFVLLTATSVAIWVWVIMVYTYGNKQHNSTPTWDDPTLATA 285
Db 743 TLVAFTRKCPENFNEAKFEGFTMYTTCIIWLAFLPIFYVTSDDYRVQT---TTMCVSVS 799

QY 286 ANAWAFVLFVIVPEVSQVTKSSPEQSYQGDWYPTRGV 323
Db 800 LSG-SVVLGCLFAPKLHILFPQKNVNVSHRAPTSRFG 836

RESULT 7

C;Keywords: differentiation; G protein-coupled receptor; receptor; transmembrane protein

F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-879/Product: metabotropic glutamate receptor subunit 3 #status predicted <MAT>
F;577-599/Domain: transmembrane #status predicted <TM1>
F;614-634/Domain: transmembrane #status predicted <TM2>
F;646-664/Domain: transmembrane #status predicted <TM3>
F;689-709/Domain: transmembrane #status predicted <TM4>
F;735-756/Domain: transmembrane #status predicted <TM5>
F;770-791/Domain: transmembrane #status predicted <TM6>
F;804-828/Domain: transmembrane #status predicted <TM7>

Query Match 6.9%; Score 160.5; DB 2; Length 879;
Best Local Similarity 22.5%; Pred. No. 3,7e-05;
Matches 69; Conservative 56; Mismatches 123; Indels 59; Gaps 14;

Qy 19 PGAWQAQGHVPPGCSQGLNPLNYNLCDR-----SGAWGVLEAVAGAGIVTTFVLTIILVA- 73
Db 551 PQQPTADL-SGC-----YNPEDYIRWEDAWAGPVTIACLGPMCTCIVITVFIKH 601

Qy 74 -SLPFPVQDTRKRSLLGTQVFELL--GTGLGFLCYFACVWKPDPFSTCASRRFLGVLFAIC 130
Db 602 NNTPLVRAS-----GRELCVILLFGVSLSCYMTFFETIAKPSVICALRRLGLGTSAIC 655

Qy 131 FSCLAHHVFAINFILAR-----KNIGPRGWVI-----FTVALLLTLEVEINTEWLIITLV 180
Db 656 YSALLTKT--NCIARIFDGVKNGAQRKFTSPSSQVFICLGLILVQIVMVSVWLILETP 712

Qy 181 RGSGEPPGQNSSAGWAVASPCALANMDFWALIVYVLLLLGAFGLGAWPALCGRYK-RWR 239
Db 713 GTRRYTLPKEKET-----VILKCNVKSMLSLIYDVVLVI-----LCTVYAKPTR 759

Qy 240 K-----HGVFVLLTATFSAIVWVIMVYTYGNKQNSPTWDDPTAIALAANAFAVL 293
Db 760 KCPENFNEAKFIGTMYTTCIIWLAFLPYFVYSSDYRVQT---TTMCISVSLSGSEFVLG 816

Qy 294 FYVIVEV 300
Db 817 CLFAPKV 823

RESULT 4
A36455
A:Title: Induction of cell fate in the Drosophila retina: the bride of sevenless protein
A:Reference number: A36455; MUID:91115074
A:Accession: A36455
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-896 <HAR>
A:Cross-references: GB:X55887
R: Hart, A.C.; Harrison, S.D.; Van Vactor Jr., D.L.; Rubin, G.M.; Zipursky, S.L.
Proc. Natl. Acad. Sci. U.S.A. 90, 5047-5051, 1993
A:Title: The interaction of bride of sevenless with sevenless is conserved between Drosophila and vertebrates
A:Reference number: A47350; MUID:93281693
A:Accession: B47550
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 4-896 <HA2>
A:Cross-references: GB:L08133
C:Genetics:
A:Gene: FlyBase:boss
A:Cross-references: FlyBase:FBgn0000206
A:Introns: 18/1; 175/1; 325/1; 811/3
C:Keywords: transmembrane protein

Query Match 6.8%; Score 159; DB 2; Length 896;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2002, 09:29:21 ; Search time 20.9 seconds
(without alignments)
2027.531 Million cell updates/sec

Title: US-09-895-686-1
Perfect score: 2326
Sequence: 1 MAHKALVMCLGLFLFLFP.....ATPKDGKNSQVFRNPYWD 441

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	167	7.2	1099	2 T16283	hypothetical prote
2	160.5	6.9	879	2 JH0562	metabotropic gluta
3	160.5	6.9	879	2 JC7160	metabotropic gluta
4	159	6.8	896	2 A36455	bride of sevenless
5	157	6.7	896	2 S26740	gene boss protein
6	152	6.5	872	2 JH0561	metabotropic gluta
7	147	6.3	999	2 T27628	hypothetical prote
8	146	6.3	893	2 A47550	bride of sevenless
9	142	6.1	1199	2 A41939	G protein-coupled
10	133	5.7	1218	2 S71376	glutamate receptor
11	132.5	5.7	551	2 T30806	metabotropic gluta
12	125	5.4	908	2 I49142	metabotropic gluta
13	118.5	5.1	1079	2 I53362	calcium/polyvalent
14	118	5.1	1085	2 S40476	Ca(2+)-sensing rec
15	117.5	5.1	699	2 H87275	thio-disulfide int
16	117.5	5.1	1078	2 A56715	calcium receptor (
17	117.5	5.1	1088	2 B56715	calcium receptor (
18	112.5	4.8	912	2 JH0563	metabotropic gluta
19	110.5	4.8	871	2 A48742	metabotropic gluta
20	108	4.6	347	2 G89003	protein T24A6.6 [i
21	108	4.6	487	2 S22350	secy protein - Hal
22	107	4.6	477	2 AB2866	exoc-like protein
23	107	4.6	480	2 H97642	exoc-like protein
24	107	4.6	646	2 AC1339	ABC transporter (p
25	106.5	4.6	477	2 D70546	hypothetical prote
26	106.5	4.6	494	2 B89827	hypothetical prote
27	106.5	4.6	642	2 H83743	ABC transporter (p
28	106	4.6	409	2 E97993	cell division prot
29	106	4.6	463	2 AC0281	probable amino aci

ALIGNMENTS

RESULT 1

T16283

hypothetical protein F35H10.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16283

R:Woessner, J.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid F35H10.

A:Reference number: Z18490

A:Accession: T16283

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-1099 <WOE>

A:Cross-references: EMBL:U40934; NID:g1072149; PID:g1072155; PIDN:AAA81683.1; CESP:F3

C:Gene: CESP-F35H10.10

A:Gene: CESP-F35H10.10

A:Introns: 35/3; 214/3; 548/1; 666/3; 740/1; 798/1; 869/2; 984/1; 1059/3

Query Match

Best Local Similarity 7.2%; Score 167; DB 2; Length 1099;

Matches 70; Conservative 55; Mismatches 140; Indels 72; Gaps 11;

QY	46	SGAWG-----IVLEAVAGAGIVTFTVLITILVASLPFVQDTKRSLLGTQVFLGLTLC	99
DB	710	TGHRDRPHNYVLLALITTVLVVVAIAVLVLVLK--YLRVVKGNOSLG--ISLLIGIII	765
QY	100	LFCLVFACVVKPDFSTCASRRFLGVLFATCFSCLAHVAFALNFKKNHGRGWIFTV	159
DB	766	LYSTAFFFPDPTSCRLRVILHGLGYTICFGVMIKAKATQLRNAETLGF-----	816
QY	160	ALLTLVEIINTEWLIITLVRGSGEGPGQGNSSAGWAV-----ASPCA	204
DB	817	---TAIHISFNWVLLFFI-----VGQIALSISWLEFPFMSTICVIDTNVQRMMCTM	867
QY	205	ANMDFVMIYVMLLIGAPLGAMPALCGR-YRRKRKHGVFVLLTTATSVAIWVWIVMY	263
DB	868	GKVEFVVSFNYFMILI---FMALFISMLNRNKRNYKETKWLILYSTVLCFFTWAVITLY	924
QY	264	TYGNKQHNSTWDDPTLATIALAANAFVLFYVIVPEVSQVTKSSP-EOSYOGDMYPTRGV	322
DB	925	LVLDHE-----FRDTVIVVELVACATILGLFGPKIYIILLSEVPVVAFKRDPFPN---	976
QY	323	GYETILKEQKQSMFVENKAFSDPEVAAKRPVSPYS	359
DB	977	-----HTDLFEKDDDLPSQRAVSPAS	997

RESULT 2

JH0562

metabotropic glutamate receptor 3 precursor - rat

```
QY 102 CLVFACVVKPDFSTCASRRRLFGVLFAICFSCLAHV--FALNFLAR-KNHGPRGW---- 154
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
663 SSSLEFFIGEPQDWTCRLRQPAFGISFVLCISILVKTNRVLLVFEAKIPTSFHRKMWGLN 722
QY 155 VIFTVALLLLTLVEVIINTEWLIITLVREGSGEGGPGQNSSAGWAVASPCAIANMDFVWALI 214
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
723 LQFLLVFLCTFMQILICIIWLY-----TAPSSYRNHELEDEII 761
QY 215 YV-----MLLLGAFLG---AWPALCGRYK-RWRK-----HGVEVLLTTTATSVAIWVW 259
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
762 FITCHEGSLMALGSLIGYTCLLAAICEFFPAKSKKLPENFNEAKFITFSMLIPIFWISF 821
QY 260 IVMY--TYG 266
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
822 IPAYASTYG 830
```

Search completed: September 16, 2002, 09:36:53
Job time: 222 sec

```
FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 461 537 MISSING (IN ISOFORM B).
FT CONFLICT 45 45 A -> S (IN REF. 2).
FT CONFLICT 304 404 L -> P (IN REF. 3).
FT CONFLICT 410 410 G -> D (IN REF. 2 AND 3).
FT CONFLICT 566 566 V -> A (IN REF. 2).
FT CONFLICT 595 595 Y -> H (IN REF. 2, 3 AND 5).
FT CONFLICT 610 610 E -> V (IN REF. 5).
FT CONFLICT 814 814 F -> L (IN REF. 5).
FT CONFLICT 889 889 L -> I (IN REF. 2).
FT CONFLICT 906 909 TGSN -> SGWI (IN REF. 2).
FT CONFLICT 1057 1057 V -> M (IN REF. 2).
FT CONFLICT 1064 1064 V -> A (IN REF. 2).
FT CONFLICT 1076 1076 I -> V (IN REF. 2).
SQ SEQUENCE 1079 AA; 120839 MW; AAF8D8D472736D6E CRC64;

Query Match 5.1%; Score 119.5; DB 1; Length 1079;
Best Local Similarity 23.7%; Pred. No. 0.4; Indels 55; Gaps 13;
Matches 59; Conservative 34; Mismatches 101;

QY 48 AW----GIVLEAVAGAGI-VTTFLVLT-I-ILVASLPFVQDTKKRSLGTVQVFFLLGTGLGF 101
Db 607 AWTEPGIALTLFAVLGIFLTAFLGVFIKFRNTPIVKATNRE----LSYLLLSLCCF 662
QY 102 CLVPACVVKPDFSTCARRFLGVFLFAICFSLAAHV--PALNFLAR-KNHGPRGW---- 154
Db 663 SSSLFFIGEPODWTCLRQPAFGISFVLCISLVKTRNVLVFEAKIPTSFHRKMWGLN 722
QY 155 VITFVALLTLVEIINTEWLLITLVIRGSGEGPGQGNSSAGWAVSCAATANDFYMALI 214
Db 723 LQELLVFLCTFMQIVICIIWLY-----TAPPSSVRNHELEDEII 761
QY 215 YV-----MLLLGAFLG---ANPACGRYK-RWRK-----HGVFVLLTTATSVAIWVW 259
Db 762 FITCHEGLMALGSLIGTYCTLLAALICFFFAFKSRKLPENFEAKFTFSMLFIFFIWIWF 821
QY 260 IVWY--TYG 266
Db 822 IPAYASTYG 830

RESULT 15
CASR_RAT STANDARD; PRT; 1079 AA.
AC P48442;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Extracellular calcium-sensing receptor precursor (CaSR) (Parathyroid
GN Cell calcium-sensing receptor).
OS CASR OR GPRC2A OR PCAR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
[1]
SEQUENCE FROM N.A.
RP STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney outer medulla;
RC MEDLINE=95116508; PubMed=7816802;
RX Riccardi D., Park J., Lee W., Gamba G., Brown E.M., Hebert S.C.;
RA "Cloning and functional expression of a rat kidney extracellular
RT calcium/polyvalent cation-sensing receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 92:131-135(1995).
RN [2]
```

```
RP SEQUENCE OF 1-294 FROM N.A.
RC STRAIN=WISTAR;
RX MEDLINE=95241465; PubMed=7724534;
RA Ruat M., Stowman A.M., Snyder S.H.;
RT "Calcium sensing receptor: molecular cloning in rat and localization
to nerve terminals."
RL Proc. Natl. Acad. Sci. U.S.A. 92:3161-3165(1995).
CC -!- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10354; AAC52149.1; -
DR EMBL; U20289; AAC52195.1; -
DR GCRDB; GCR_1449; -
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF000003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PRINTS; PR00248; GPCRMR.
DR PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; 1.
DR PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; 1.
DR PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; 1.
DR PROSITE; PS00982; G_PROTEIN_RECEP_F3_4; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1079 EXTRACELLULAR CALCIUM-SENSING RECEPTOR.
FT DOMAIN 20 612 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 613 635 I (POTENTIAL).
FT DOMAIN 636 649 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 650 670 II (POTENTIAL).
FT DOMAIN 671 681 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 682 700 III (POTENTIAL).
FT DOMAIN 701 724 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 725 745 IV (POTENTIAL).
FT DOMAIN 746 769 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 770 792 V (POTENTIAL).
FT DOMAIN 793 805 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 806 828 VI (POTENTIAL).
FT DOMAIN 829 836 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 837 862 VII (POTENTIAL).
FT DOMAIN 863 1079 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1079 AA; 120867 MW; D7664550361F9736 CRC64;
```

```
Query Match 5.1%; Score 118.5; DB 1; Length 1079;
Best Local Similarity 23.7%; Pred. No. 0.47;
Matches 59; Conservative 34; Mismatches 101; Indels 55; Gaps 13;

QY 48 AW----GIVLEAVAGAGI-VTTFLVLT-I-ILVASLPFVQDTKKRSLGTVQVFFLLGTGLGF 101
Db 607 AWTEPGIALTLFAVLGIFLTAFLGVFIKFRNTPIVKATNRE----LSYLLLSLCCF 662
```


Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 "A family of metabotropic glutamate receptors.";
 Neuron 8:169-179(1992).
 (4)
 ALTERNATIVE SPLICING (ISOFORM 1C).
 TISSUE=Brain;
 MEDLINE=93066232; PubMed=1438218;
 Pin J.-P., Waeber C., Prezeau L., Bockaert J., Heinemann S.F.;
 "Alternative splicing generates metabotropic glutamate receptors
 inducing different patterns of calcium release in Xenopus oocytes.";
 Proc. Natl. Acad. Sci. U.S.A. 89:10331-10335(1992).
 CC -!- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
 MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-
 CALCIUM SECOND MESSENGER SYSTEM. MAY PARTICIPATE IN THE CENTRAL
 ACTION OF GLUTAMATE IN THE CNS. SUCH AS LONG-TERM POTENTIATION IN
 THE HIPPOCAMPUS AND LONG-TERM DEPRESSION IN THE CEREBELLUM.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1A (SHOWN HERE), 1B AND 1C;
 ARE PRODUCED BY ALTERNATIVE SPLICING. 1B AND 1C ARE C-TERMINALLY
 TRUNCATED FORMS OF 1A.
 CC -!- TISSUE SPECIFICITY: IS PREDOMINANTLY EXPRESSED IN CEREBELLAR
 PURKINJE CELLS, CA2-CA3 PYRAMIDAL CELLS OF THE HIPPOCAMPUS, AND
 MITRAL AND TUFTED CELLS OF THE OLFACTORY BULB.
 CC -!- MISCELLANEOUS: ACTIVATED BY QUISQUALATE > GLUTAMATE > IBOTENATE >
 TRANS-1-AMINOCYCLOPENTYL-1,3-DICARBOXYLATE; INHIBITED BY
 2-AMINO-3-PHOSPHONOPROPIONATE.
 CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR8.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X57569; AAA40799.1; -;
 DR EMBL; M61099; AAL19497.1; -;
 DR EMBL; S48085; AAB24138.1; -;
 DR PIR; S15362; S15362.
 DR PIR; A41939; A41939.
 DR GCRdb; GCR_0216; -;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam; PF00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMRG.
 DR PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECEP_F3_4; 1.
 DR PROSITE; PS02599; G_PROTEIN_RECEP_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 Multigene family; Alternative splicing.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 1199 METABOTROPIC GLUTAMATE RECEPTOR 1.
 FT CHAIN 19 592 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 593 615 I (POTENTIAL).
 FT DOMAIN 616 629 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 630 650 II (POTENTIAL).
 FT DOMAIN 651 661 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 662 680 III (POTENTIAL).
 FT DOMAIN 681 706 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 707 727 IV (POTENTIAL).
 FT DOMAIN 728 750 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 751 772 V (POTENTIAL).
 FT DOMAIN 773 785 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 786 808 VI (POTENTIAL).
 FT DOMAIN 809 814 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 815 840 VII (POTENTIAL).
 FT DOMAIN 841 1199 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 1014 1034 GLN/PRO-RICH.
 FT DOMAIN 1074 1080 GLN/PRO-RICH.

FT DOMAIN 1126 1135 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 1140 1199 SER-RICH.
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 887 906 NSNGKSVSWSEPGGRQAPKG -> KKRQPEFSPSQPSAH
 AOL (IN ISOFORM 1B).
 FT VARSPLIC 907 1199 MISSING (IN ISOFORM 1B).
 FT VARSPLIC 888 897 SNKGSVSWSE -> FALDRQNTVY (IN ISOFORM 1C).
 FT VARSPLIC 898 1199 MISSING (IN ISOFORM 1C).
 SQ SEQUENCE 1199 AA; 133235 MW; EEE5A04C50694B9F CRC64;
 Query Match 6.1%; Score 142; DB 1; Length 1199;
 Best Local Similarity 22.4%; Pred. No. 0.011;
 Matches 85; Conservative 54; Mismatches 153; Indels 88; Gaps 19;
 QY 20 GWAQGHVPPGCCQGLNPLVYNLCDSGANGIVLEAVAGAGIVTTFVLTILV--ASLPF 77
 DB 568 GWPNAEL-----TGCEPIPVRYLEWSDIESIIAIFSCGLILVTLFVTLFVLYRTPV 622
 QY 78 VQDTRKRLSLGTQVFFLLGT-LGLFCLVFACVVKVPDFSTCASRRFLFGVLFACFSCIAA 136
 DB 623 VKSSRELC-----YIILAGIFLGVC-PFTLIAKPTTSCYLQRLVGLSSAMCYSAL-- 675
 QY 137 HVFALNPLARKNIG-----PR---GWVIFTVALLLTLVEVINTWELIITLVGRSG 184
 DB 676 -VTKTRRIARILAGSKKKICTRKPFRMSAWAQVILISIVQLT-----LVVTLI---- 725
 QY 185 EGGPGQGNSSAGWAVASP-----CATANMDFVMAIIVMVLILLGLAFLGAW 228
 DB 726 -----IMPEPMPILSYPSIKEVYLICNTSNLGVAPVGVNGLLINSCTIYAF 772
 QY 229 -----PALCGRYKRWKRGHVLLTTATSVAIWVWVIVMYTYGNKHNSPTDPTLAI 283
 DB 773 KTRNVPAFNEAK-----YIAFTMYTTCIIWLAFVPIY-FGSNYKIITCFVAVLSVT 824
 QY 284 LAANAFVLFYVI---PE---VSQVTKSSPEQSYQGD-MYPTRGVGYETILKEQK-GQS 335
 DB 825 VALGCMFTPKMYIIIAKPERNVRSAFTTSDVRMHVGDGKLPKCRSNFTLNFIRKKPKG 884
 QY 336 MEVEN-KAFSDPBPVAAKRP 354
 DB 885 NANSNGKSVSWSEPGGRQAP 904
 RESULT 11
 MGR8_MOUSE
 ID MGR8_MOUSE STANDARD; PRT; 908 AA.
 AC P47743;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 8 precursor.
 GN GRM8 OR GPRC1H OR MGLUR8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Retina;
 RX MEDLINE=95239344; PubMed=7722646;
 RA Duvoisin R.M., Zhang C., Ramonell K.;
 RT "A novel metabotropic glutamate receptor expressed in the retina and
 RT olfactory bulb.";
 RL J. Neurosci. 15:3075-3083(1995).
 CC -!- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 CC ACTIVITY.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN OLFACTORY BULB,

```

Db 549 SLTGCFELPQXYIRWGDAMAVGPVTIAC-----LGA-LAT 582
QY 64 TFVLITILVASLDFVQDTRKRSLLGTQVFFELTGL-FCLVFACVVKPDFSCASRRFL 122
Db 583 LFLVGVFVRHNATPVVKASGREL-----CYILGGVFLCYCMTIFIAKPESTAVCTIRRLG 638
QY 123 FGVLFACFSCAAHFAVNLARKNHGPRGW-----IFTVALLLTLV--EVIINT 172
Db 639 LGTAFSVCYSAALLTKT---NRIARIFGAREGAORPREFISPAQVAICLALISGQLLIV 695
QY 173 EWLIIILVRSGEGGQGNSSAGWASVCAIANMDFVMAIYVMLLLLGAFGLGAWPALC 232
Db 696 AMLVLRAPGTGKETAPERRE-----VTLRCNHRDASHMLGSLAYNLLI-----ALC 742
QY 233 GRYK-RWRK-----HGVFVLLTTATSVAIWVIMVMTYGNKQHNPSPTWDDPTLAIALA 285
Db 743 TLAFNTRKPCPENFAKFIGFTMYTTCILWALLPIFYVTSDDYRQV---TTMCVSVS 799
QY 286 ANAWAFVLFVIVPEVSQVTKSSPEQSQGDMDYPTRGVG 323
Db 800 LSG-SVVLGCLFAPKLHIILFQPKNVVSHRAPTSRFG 836

RESULT 9
BOSS_DROVI STANDARD; PRT; 893 AA.
AC Q24738;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Bride of sevenless protein precursor.
GN BOSS.
OS Drosophila virilis (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7244;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93281693; PubMed=8506350;
RA Hart A.C., Harrison S.D., van Vactor D.L. Jr., Rubin G.M.,
RA Zipursky S.L.;
RT "The interaction of bride of sevenless with sevenless is conserved
RL between Drosophila virilis and Drosophila melanogaster.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:5047-5051(1993).
CC -1- FUNCTION: ACTS AS A LIGAND FOR SEVENLESS TYROSINE-KINASE RECEPTOR
CC DURING EYE DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L08132; AAA72332.1;
DR FlyBase; FBgn0013105; Dvir\boss.
DR InterPro: IPR002956; Bride_of_7less.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR PRINTS; PR01223; BRIDOF7LESS.
DR Transmembrane; Glycoprotein; Vision; Signal.
KW SIGNAL
FT CHAIN 1 30 POTENTIAL.
FT DOMAIN 31 893 BRIDE OF SEVENLESS PROTEIN.
FT TRANSMEM 32 530 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 531 551 POTENTIAL.
FT TRANSMEM 563 583 POTENTIAL.
FT TRANSMEM 607 627 POTENTIAL.
FT TRANSMEM 630 650 POTENTIAL.
FT TRANSMEM 653 673 POTENTIAL.
FT TRANSMEM 692 712 POTENTIAL.

```

```

FT TRANSMEM 722 742 POTENTIAL.
FT TRANSMEM 772 772 POTENTIAL.
FT DOMAIN 773 893 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 893 AA; 99939 MW; ABB9DAC21DB680C3 CRC64;

Query Match 6.3%; Score 146; DB 1; Length 893;
Best Local Similarity 20.8%; Pred. No. 0.0043;
Matches 71; Conservative 58; Mismatches 136; Indels 76; Gaps 15;

QY 34 GLNPLYNCLDRSGA--WGIVLEAVAGAGIVTTFVLITILVASLFPVQDTKKRSLLG--- 88
Db 503 GENRRPFFLDGESVWFWRKLDTWATGLTAAILGLIATLAILVFI---VVRISLGDVF 559
QY 89 -----TQVFFELGLTGLFC--LVFAC-----VVKPDFST---CASRFLFCVL 127
Db 560 EGNPVTISILLLSLILVFCSEVFPFSEYVGEQRNSHVTFEDVHTLNTLCGVRVFIITLVY 619
QY 128 AICFSCIAHVFAL-----NFLARKNHGPRGWITFTVALLTLVEVIINTWLLITLV 180
Db 620 CFVFSLLLCRAVMLASTSGSEGGFLSHVN---GYIOAIIICVLSVFQVQMSVQLLV--M 673
QY 181 RGSSEGGPOGNSAG--WAVASPCAIAIMDFVMAIYVMLLLLGAFGLGAWPALCGRYKRW 239
Db 674 HLASESVSCENIYYGRW-----LWGLLAYDFLLLCSLVSLVFPFYRSQRNRY 720
QY 240 KHGVFVLLTTATSVAIWVIMVMTYGNKQHNPSPTWDDPTLAIALAANAFVLFVVP- 298
Db 721 E-GILIVIGAVLIIINSWIAMSMEGDE-----WRDAIPLGMQASGAWVLVGLILPR 773
QY 299 -----EVSQVTKSSPEQS-----YQGDMDYPTRGVGYETI 327
Db 774 TFLIVRGIERSDIAQALPSLTSLAFAQNNOYSSEQSVYECV 814

RESULT 10
MGRL_RAT STANDARD; PRT; 1199 AA.
AC P23385;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metabotropic glutamate receptor 1 precursor.
GN GRM1 OR GPRC1A OR MGLUR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=91156047; PubMed=1847995;
RA Masu M., Tanabe Y., Tsuchida K., Shigemoto R., Nakanishi S.;
RT "Sequence and expression of a metabotropic glutamate receptor.";
RL Nature 349:760-765(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=92022526; PubMed=1656524;
RA Houamed K.M., Kuijper J.L., Gilbert T.L., Haldeman B.A., O'Hara P.J.,
RA Mulvihill E.R., Almers W., Hagen F.S.;
RT "Cloning, expression, and gene structure of a G protein-coupled
RL glutamate receptor from rat brain.";
RL Science 252:1318-1321(1991).
RN [3]
RP ALTERNATIVE SPLICING (ISOFORM 1B).
RX TISSUE=Brain;
RX MEDLINE=92110002; PubMed=1309649;

```

```

NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=95346007; PubMed=7620613;
Flor P.J., Lindauer K., Puttner I., Ruegg D., Lukic S., Knopfel T.,
Kuhn R.;
"Molecular cloning, functional expression and pharmacological
characterization of the human metabotropic glutamate receptor type
2.";
Eur. J. Neurosci. 7:622-629(1995).
CC -I- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
CC MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.
CC MAY MEDiate SUPPRESSION OF NEUROTRANSMISSION OR MAY BE INVOLVED IN
CC SYNAPTONEGENESIS OR SYNAPTIC STABILIZATION.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- TISSUE SPECIFICITY: WIDELY EXPRESSED IN DIFFERENT REGIONS OF THE
CC ADULT BRAIN AS WELL AS IN FETAL BRAIN.
CC -I- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR3.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licenssee@isb-sib.ch).
-----
CC EMBL; L35318; AAA76855.1; -.
CC GCRdb; GCR_1846; -.
CC MIM; 604099; -.
CC InterPro; IPR001182; ANF_receptor.
CC InterPro; IPR000337; GPCR_Mgr.
CC Pfam; PF000003; 7tm_3; 1.
CC Pfam; PF01094; ANF_receptor; 1.
CC PRINTS; PR00248; GPCRMGR.
CC DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
CC DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
CC DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
CC DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
CC DR PROSITE; PS0259; G_PROTEIN_RECP_F3_4; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family.
FT SIGNAL 1 18
FT CHAIN 19 872 METABOTROPIC GLUTAMATE RECEPTOR 2.
FT DOMAIN 19 567 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 568 590 I (POTENTIAL).
FT DOMAIN 591 604 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 605 625 II (POTENTIAL).
FT DOMAIN 626 636 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 637 655 III (POTENTIAL).
FT DOMAIN 656 679 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 680 700 IV (POTENTIAL).
FT DOMAIN 701 725 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 726 747 V (POTENTIAL).
FT DOMAIN 748 760 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 761 783 VI (POTENTIAL).
FT DOMAIN 784 793 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 794 819 VII (POTENTIAL).
FT DOMAIN 820 872 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 547 547 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 872 AA; 95507 MW; 058608C35C701E9D CRC64;

Query Match 6.3%; Score 146; DB 1; Length 872;
Best Local Similarity 21.9%; Pred No. 0.0042;
Matches 74; Conservative 49; Mismatches 145; Indels 70; Gaps 14;

QY 6 ALVMCLGIP--LFLFPGAWAQGHHVPPCGSGNLPLYNLCDRSGAWGVILEAVAGIVT 63
:| | | | : ||| | | |

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR	EMBL: Z47073: CAA87374.1; -	
DR	WormPep; ZC506.4; CE01682.	
DR	InterPro: IPR001828; ANF_receptor.	
DR	InterPro: IPR000337; GPCR_Mgr.	
DR	Pfam: PF00003; 7tm3; 1.	
DR	Pfam: PF01094; ANF_receptor; 1.	
DR	PRINTS: PR00248; GPCRMRG.	
DR	PROSITE; PS00979; G_PROTEIN_RECEP_F3_1; 1.	
DR	PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; 1.	
DR	PROSITE; PS00981; G_PROTEIN_RECEP_F3_3; 1.	
DR	PROSITE; PS0259; G_PROTEIN_RECEP_F3_4; 1.	
KW	G-protein coupled receptor; Transmembrane; Glycoprotein.	
FT	TRANSMEM 682 704	POTENTIAL.
FT	TRANSMEM 719 739	POTENTIAL.
FT	TRANSMEM 751 769	POTENTIAL.
FT	TRANSMEM 792 812	POTENTIAL.
FT	TRANSMEM 836 857	POTENTIAL.
FT	TRANSMEM 871 893	POTENTIAL.
FT	TRANSMEM 904 929	POTENTIAL.
FT	DOMAIN 888 891	POLY-PHE.
FT	CARBOHYD 518 518	N-LINKED (GLCNAC..) (POTENTIAL).
SQ	SEQUENCE 999 AA; 1133275 MW; AB22AF2A28D9ADCF CRG64;	

Query Match	6.3%	Score 147;	DB 1;	Length 999;
Best Local Similarity	20.1%	Pred. No. 0.0041;		
Matches 85;	Conservative	61;	Mismatches 143;	Gaps 20;
			Indels 134;	

Qy	30	GC	SQGLNPLYNLCDRSGAWGIVLE-----AVAGAGIVTTFVLTITL 71
		:	: : : :
Db	650	CSEQG----	-WMPTKDRKGCFDLSLSQLKYMRWRSMYSLVPTILAVFGI-IATLFVIVVV 704
Qy	72	VASLPFVD	TKKSLTGCTQVFLLGTGLGFCLFWACWKPFDFSTCASRRFLFCVLFAICF 131
Db	705	IYNETPVKASGREL----	-SVILLISIMCYCMTFFVLLSRPSAIVCAIKR--TCIGFA--F 757
		:	: : : :
Qy	132	SCIAAHVFALN-----	FLARKNHGRP-----GWIFTVALLLTVEVIINTEWLIIITLV 180
		:	: : : :
Db	758	SCLYSAMFVKTNRIERIFSRSAQRPRFISPISQVYMTA--MLAGVOLIGSLIWSLV-- 813	
Qy	181	RSGEGGPQGNSSAGW-----	-AVASPCAIANMDFVALIYVMULLLGAFLGAWPALC 232
Db	814	-----PPGWRHHYPTRDQVVLTCNVDPDHFLXSLAYDGFLV-----LC 852	
Qy	233	GRYK-RWRK-----	HGVFLLTTATSVAIWVWIYMYTGNKHNSPTWDPTDLTALA 285
		:	: : : :
Db	853	TTVAVKRKVPENETKFICFSMYTCVVWLSWIEFF-FGCGSDFOIQTSLICISMS 911	
Qy	286	ANAWAFVLYVIPESQVTKSSPE-----	-OSYOGDMYPTRGVGYETILKEQGQSFMFE 339
		:	: : : :
Db	912	ANVALACIF-----	SPKLWIIIPKHKH-----VRKQBGESMLNK 946
Qy	340	NK-----	-AFSMDPEVAAKRVPSPSYGNGOLLTSVYQPTEMALMHKHPSSGAYD 388
		: : :	: : : :
Db	947	SSRLGNCSSRLCANSIDEP-----	-NQYTALLDTSRRRSRKKT SQPTSTSSAHD 996
Qy	389	IIL 391	

RESULT 7

	MGR1_HUMAN	STANDARD;	PRT: 1194 AA.
ID	MGR1_HUMAN	STANDARD;	PRT: 1194 AA.
AC	Q13255; Q13256; Q14757; Q14758;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Metabotropic glutamate receptor 1 precursor.		
GN	GRM1 OR GPRC1A OR MGLU1R1.		


```

QY 123 FGVLEAIFCSLAHVAFLNFKNHGPRGW-----IFTVALLTLV--EVIINT 172
Db LGTAFSCVYSALLTKT---NRIATFGGAREGAQRFISPAQVAICLALISSQLLIVA 695
QY 173 EWLIITLVROSGGGGQGNSSAGVAPCAIANMDFVMAIIVYVMLLLGAFLGAWPALC 232
Db 696 AYLWEAPGPGKETAPERRE---VVTLRNHRDASMLGSLAYNVLLI-----ALC 742
QY 233 GRYK-RWRK-----HCVFVLLTATSAVAIVVWVYTYGKOHNSPTWDDPTLALALA 285
Db 743 TLVAFTRKCPENFNEAKFGFTMYTTCIIWLAFLPIFYVYVTSDDYRVQT---TTMCSVSVS 799
QY 286 ANAWAFVLFVIVPEVSVOTKSSPEQSYQGDMPYTRGVG 323
Db 800 LSG-SVVLGCLFAPKLHILFQPKNVVSHRAPTSREG 836

RESULT 5
MGR3_HUMAN
ID MGR3_HUMAN STANDARD; PRT; 877 AA.
AC Q14832;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metabotropic glutamate receptor 3 precursor.
GN GRM3 OR GPRC1C OR MGLUR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96437205; PubMed=8840013;
RA Makoff A., Volpe F., Lelchuk R., Harrington K., Emson P.;
RT "Molecular characterization and localization of human metabotropic
RT glutamate receptor type 3.";
RL Brain Res. Mol. Brain Res. 40:55-63(1996).
CC -!- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
CC ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR2.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X77748; CAAS4796.1; -.
DR GCRDB; GCR_2070; -.
DR MTM; 601115; -.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PRINTS; PR00248; GPCRMRG.
DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
DR PROSITE; PS0259; G_PROTEIN_RECP_F3_4; 1.
DR KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 877 METABOTROPIC GLUTAMATE RECEPTOR 3.
FT DOMAIN 21 574 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 575 597 I (POTENTIAL).
FT DOMAIN 598 611 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 612 632 II (POTENTIAL).

```

```

FT DOMAIN 633 643 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 644 662 III (POTENTIAL).
FT DOMAIN 663 686 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 687 707 IV (POTENTIAL).
FT DOMAIN 708 732 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 733 754 V (POTENTIAL).
FT DOMAIN 755 767 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 768 790 VI (POTENTIAL).
FT DOMAIN 791 800 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 801 826 VII (POTENTIAL).
FT DOMAIN 827 877 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 877 AA; 98619 MW; 66F28663CE35F740 CRC64;

Query Match 6.3%; Score 147; DB 1; Length 877;
Best Local Similarity 21.8%; Pred. No. 0.0036;
Matches 61; Conservative 55; Mismatches 106; Indels 58; Gaps 12;

QY 48 AWGIVLEAVAGAGIVTFTVLIILVA--SLPFDVQDTKRSLLTQVFLGLTGL-FCLV 104
Db 573 AWAIGPVTIACLGFMCTCMVTVFIKHNTPLVKASGRE-----LCYILLFGVGLSYCMT 627
QY 105 FACVVRPDFSTCASRRFLGVLEAIFCSLAHVAFLNFKL-----KNHGPRGWVI--- 156
Db 628 FFFIAKPSVPICALRLRLGLSGSAICYSAULTKT---NCIARIFDGVKNGAQRKPFISPS 684
QY 157 --FTVALLTLVEVINTETMLIITLVRSGEQGGPQGNSSAGWAVAS-----PCAIANM 207
Db 685 SQVFICGLGLILVQIVVSVMLIL-----EAPCTRRYTLAEKRETVILKCNVKDS 733
QY 208 DFVMAIYVMLLLGLFAGLWALPGQRYK-RWRK-----HGTVFLTLTATSAIVWVWI 260
Db 734 SMLISLTYDVLVI-----LCTVVAEKTRKCPENFNEAKFIGTMTVTTCTIWLAF 784
QY 261 VMITYGNKHNSPTWDDPTLALANAWAFVLFYVIPEV 300
Db 785 PIFVYTSDDYRVQT---TTMCISVSUGFVVLGCLFAPKV 821

RESULT 6
MGR1_CAEL
ID MGR1_CAEL STANDARD; PRT; 999 AA.
AC Q09630;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Probable metabotropic glutamate receptor mgl-1.
GN MGL-1 OR ZC506.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Harris B.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
CC ACTIVITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR3.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```


[illegible]

G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 Multigene family.
 KW SIGNAL 1 22 POTENTIAL
 FT CHAIN 23 879 METABOTROPIC GLUTAMATE RECEPTOR 3.
 FT DOMAIN 23 576 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 577 599 I (POTENTIAL).
 FT DOMAIN 600 613 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 614 634 II (POTENTIAL).
 FT DOMAIN 635 645 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 646 664 III (POTENTIAL).
 FT DOMAIN 665 688 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 689 709 IV (POTENTIAL).
 FT DOMAIN 710 734 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 735 756 V (POTENTIAL).
 FT DOMAIN 757 769 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 770 792 VI (POTENTIAL).
 FT DOMAIN 793 802 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 803 828 VII (POTENTIAL).
 FT DOMAIN 829 879 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 879 AA; 98959 MW; 3E5965EDD56DEB CRC64;
 Query Match 6.9%; Score 160.5; DB 1; Length 879;
 Best Local Similarity 22.5%; Pred. No. 0.0004;
 Matches 69; Conservative 56; Mismatches 123; Indels 59; Gaps 14;
 QY 19 PGWAQGHVPPGSGQLNPLYYNLCDR-----SGAWGIVLEAVAGAGIVTFVLTIIILVA- 73
 Db 551 PGQWPTADL-SGC-----YNLPEYIKWEDAWAGPVYTIACGLGFLCTCIVITFIKH 601
 QY 74 -SLPFDVDTKRSLLGQVFL--GTLGLFCLVACVVKPDFSTCASRRFLFGVLPFAIC 130
 Db 602 NNTPLVKAS-----GRELCYILLFGVSLSYCMYTFEFAKPSVPICALRRLGLGTSFAIC 655
 QY 131 FSLAAHVFAFNFLAR-----KNHGPRGWI-----FTVALLTLTVEINTENLITLV 180
 Db 656 YSALLTKT---NCIARIFDGVKGAQPKFISPSQVFCILGLLIVGVWVSVWLILETP 712
 QY 181 RGSGEQPGQSGAGWAVASPCAIANMDFVYALIVYMLLLIGAFLGAWPALCGRYK-RWR 239
 Db 713 GTRRYTLPERRET----VILKCNKVDSSMLISLYDVLVI-----LCITYAPKTR 759
 QY 240 K-----HGVFVLLTATSVAVIIVWVIMVYTKNQKINSFTWDDPTLAIALANAFVL 293
 Db 760 KCPENFNEAKFIGFTWYTTCTIWLAFIPFYVTSSDYRVQT---TTMCISVSLSGFVVLG 816
 QY 294 FYVIVEV 300
 Db 817 CLFAPKV 823
 RESULT 3
 ID BOSS_DROME STANDARD; PRT; 896 AA.
 AC P22815; Q9VBJ5;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bride of sevenless protein precursor.
 GN BOSS OR CG8285.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91115074; PubMed=2276620;
 RA Hart A.C., Kraemer H., van Vactor D.L. Jr., Paidhungat M.,

Zipursky S.L.;
 "Induction of cell fate in the Drosophila retina: the bride of
 sevenless protein is predicted to contain a large extracellular
 domain and seven transmembrane segments.";
 Genes Dev. 4:1835-1847(1990).
 [2]
 RP REVISIONS.
 RA Kraemer H.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Brokstein P., Brotter P.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclev J.M.,
 Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of Drosophila melanogaster.";
 Science 287:2185-2195(2000).
 [4]
 RP FUNCTION.
 RX MEDLINE=91312442; PubMed=1857416;
 RA Kraemer H., Cagan R.L., Zipursky S.L.;
 RT "Interaction of bride of sevenless membrane-bound ligand and the
 sevenless tyrosine-kinase receptor.";
 RL Nature 352:207-212(1991).
 CC -!- FUNCTION: ACTS AS A LIGAND FOR SEVENLESS TYROSINE-KINASE RECEPTOR
 CC DURING EYE DEVELOPMENT.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY BY R8 PHOTORECEPTOR
 CC CELLS AND IS INTERNALIZED IN A SEV-DEPENDENT MANNER BY R7 CELLS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X55887; CAA39373.1; -.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2002, 09:33:11 ; Search time 13.4 Seconds
(without alignments)
1274.278 Million cell updates/sec

Title: US-09-895-686-1
Perfect score: 2326
Sequence: 1 MAIHKALYMCIGLPLFLPG.....ATPPKDGKNSQVFRNPYVD 441

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	174.5	7.5	976	1 MGR_DROME	P91685 drosophila
2	160.5	6.9	879	1 MGR3_RAT	P31422 rattus norv
3	155	6.7	896	1 BOSS_DROME	P22815 drosophila
4	152	6.5	872	1 MGR2_RAT	P31421 rattus norv
5	147	6.3	877	1 MGR3_HUMAN	Q14832 homo sapien
6	147	6.3	999	1 MGR1_CAEEL	Q09630 caenorhabdi
7	147	6.3	1194	1 MGR1_HUMAN	Q13255 homo sapien
8	146	6.3	872	1 MGR2_HUMAN	Q14416 homo sapien
9	146	6.3	893	1 BOSS_DROVI	Q24738 drosophila
10	142	6.1	1199	1 MGR1_RAT	P23385 rattus norv
11	125	5.4	908	1 MGR8_MOUSE	P47743 mus musculus
12	122	5.2	908	1 MGR8_HUMAN	O00222 homo sapien
13	122	5.2	908	1 MGR8_RAT	P70579 rattus norv
14	119.5	5.1	1079	1 CASR_MOUSE	Q9qy96 mus musculus
15	118.5	5.1	1079	1 CASR_RAT	P48442 rattus norv
16	118	5.1	1085	1 CASR_BOVIN	P35384 bos taurus
17	117.5	5.1	1078	1 CASR_HUMAN	P41180 homo sapien
18	116.5	5.0	877	1 MGR6_HUMAN	O15303 homo sapien
19	112.5	4.8	912	1 MGR4_RAT	P31423 rattus norv
20	111.5	4.8	912	1 MGR4_HUMAN	Q14833 homo sapien
21	110.5	4.8	871	1 MGR6_RAT	P35349 rattus norv
22	110	4.7	1065	1 EMBB_MYCAV	P71496 mycobacteri
23	108	4.6	487	1 SECK_HALMA	P28542 haloarcula
24	106	4.6	389	1 UR2R_HUMAN	Q9ukp6 homo sapien
25	104.5	4.5	2254	1 CCAG_RAT	O54898 rattus norv
26	104	4.5	1083	1 EMBB_MYCLE	Q9cda9 mycobacteri
27	103.5	4.4	2353	1 CCAH_HUMAN	Q95180 homo sapien
28	103	4.4	416	1 RHLE_PANTR	Q28812 pan troglod
29	103	4.4	567	1 YFZ1_SCHPO	Q9us40 schizosacch
30	102	4.4	1835	1 CCAL_RAT	Q9z0y8 rattus norv
31	102	4.4	2377	1 CCAG_HUMAN	O43497 homo sapien
32	101.5	4.4	612	1 ELFI_MOUSE	O60775 mus musculus
33	101	4.3	1074	1 EMBB_MYCSM	O50393 mycobacteri

RESULT 1

ID	MGR_DROME	STANDARD;	PRT;	976 AA.
AC	P91685; Q9V485;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	01-MAR-2002 (Rel. 41, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Metabotropic glutamate receptor precursor.			
GN	GLU-RA OR GLURA OR CG11144.			
OS	Drosophila melanogaster (fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RP	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RX	MEDLINE=96421661; PubMed=8824309;			
RA	Parmentier M.L., Pin J.P., Bockaert J., Grau Y.;			
RT	"Cloning and functional expression of a Drosophila metabotropic			
RT	glutamate receptor expressed in the embryonic CNS. ";			
RL	J. Neurosci. 16:6687-6694(1996).			
RN	[2]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Anatrides P.G., Scher S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	George R.A., Lewis S.E., Zhang Q., Chen L.X.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang M., Pfeiffer B.D.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Miklos G.L.G.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Baldwin D.,			
RA	Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jalali M., Kallush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,			
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,			
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,			
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,			
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,			

ALIGNMENTS

Search completed: September 16, 2002, 09:36:33
Job time: 222 sec


```
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RP TISSUE=BRIN;
RC MEDLINE=20145471; PubMed=10679242;
RX Inoue S., Sano H., Ohka M.;
RA "Growth suppression of Escherichia coli by induction of expression of
RT mammalian genes with transmembrane or ATPase domains.";
RL Biochem. Biophys. Res. Commun. 268:553-561(2000).
DR EMBL; AB030197; BAA92760.1; -.
DR MGD; MGI:1927596; Gprc5b.
FT NON_TER 1
SQ SEQUENCE 262 AA; 29305 MW; 6407A0D9FADF75F8 CRC64;

Query Match 18.3%; Score 425.5; DB 11; Length 262;
Best Local Similarity 38.5%; Pred. No. 1.2e-27;
Matches 102; Conservative 38; Mismatches 72; Indels 53; Gaps 8;

Qy 144 LARKNHGPRGVVFTVALLTLVEVIINTEWLIITLVRSGEQGGPQGNSSAGWAVASPCA 203
Db 2 LVROGTSPASQVSLALCLMLQVVIATLWLVTLRDT-----KPCA 46

Qy 204 IANMDFVALLIYVMLLLGAFGLGAWPALCGYKRWKRGVFLVLTATTSVAIVVWIVMY 263
Db 47 YEPDMFVALLIYDVLLAITLAQSLFTLCGKFKRWKVGAFILVTTFLSALIYVWWTMY 106

Qy 264 TYGN--KOHNSPTWDDPTLAIALAANAWAFVLFVIVPEVSQVTKSSPEQSQGDMYPT- 319
Db 107 LFGSLIKQGG--WSDPTLAITLAAGSGWVFVIFHAPEI-HYTLPLQENPNFYDTS 163

Qy 320 ----RGVGYETILKEQKQSGMFVENKAFSMDPEVAAKRPVSPYSGYNGQL----- 365
Db 164 QPRMRETAFD--EEMHLPAYMENKAFSMDENHAURSAGFS--NGSLQRSSSLGKK 218

Qy 366 -----LTSVQPTMALM 378
Db 219 PSSLGNRPSPAPRSNVYQPTMAVV 243

RESULT 12
Q20073 PRELIMINARY; PRT; 1404 AA.
ID Q20073; Q01363;
AC Q20073; Q01363;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 157.9 KDA PROTEIN (CELF35-1).
GN F35H10.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Woessner J.;
RT "The sequence of C. elegans cosmid F35H10.";
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
```

```

RN SEQUENCE FROM N.A.
RP Abe T., Tanemoto M., Nishio T., Hebert S.C.;
RT "Metabotropic glutamate-like sequence in C. elegans.";
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U00934; AAA81683.2; -.
DR EMBL; U00870; AAB51198.1; -.
DR InterPro; IPR000337; GPCR_Mgr.
DR InterPro; IPR000847; HTH_LysR.
DR Pfam; PF00003; 7tm_3; 1.
DR PRINTS; PR00248; GPCRMR.
DR PROSITE; PS0259; G_PROTEIN_RECEP_F3_4; 1.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; UNKNOWNW_1.
KW Hypothetical protein.
SQ SEQUENCE 1404 AA; 157943 MW; 46D8282F7FC464A0 CRC64;

Query Match 7.2%; Score 167; DB 5; Length 1404;
Best Local Similarity 20.8%; Pred. No. 3.5e-05;
Matches 70; Conservative 55; Mismatches 140; Indels 72; Gaps 11;

Qy 46 SGWG-----IVLEAVAGAGIVTFTVLTITLVASLPFVODTKKRSLLTQVFPFLGLTGL 99
Db 1015 TGHWRDHPHNVLLALITLVVVAIAVLVLVVKL--YLRVVKGNQSLG--ISLLIGIII 1070

Qy 100 LFCLVFACVVKPDFSCASRRFLFGLFAICFSCLAARHVFALNFARKNHGPRGWVFTV 159
Db 1071 LYSTAFFEVEDPTDSVCRRLVHLGLGVTICFGVMIKATOLRNAETLGF----- 1121

Qy 160 ALLTLVEVIINTEWLIITLVRSGEQGGPQGNSSAGWAV-----ASPCA 204
Db 1122 ----TAIHISFNWYMLLFFI-----VGQIALSISFLEPFMFSTIGVIDTNVQRMCTM 1172

Qy 205 ANMDFVALLIYVMLLLGAFGLGAWPALCGR-YKRWKRGVFLVLTATTSVAIVVWIVMY 263
Db 1173 GKVEFVSNFYVMIIL---FMALFISMLNRNKRNYKTKMLLYSTVLCFTTWWAVITLY 1229

Qy 264 TYGNKQNSPTWDDPTLAIALAANAWAFVLFVIVPEVSQVTKSSP-EQSQGDMYPTRGV 322
Db 1230 LVLDHE-----FRDTVIVVELVACATILGLFGPKVILLSYEPVVVAVPKRDPFN--- 1281

Qy 323 GYETILKEQKQSGMFVENKAFSMDPEVAAKRPVSPYS 359
Db 1282 -----HTDLFEKDDLPORAVSPAS 1302

RESULT 13
Q20073 PRELIMINARY; PRT; 879 AA.
ID Q20073; Q09052
AC Q20073; Q09052;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE METABOTROPIC GLUTAMATE RECEPTOR 3 PROTEIN.
GN GRM3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=20012997; PubMed=10544282;
RA Minoshima T., Nakanishi S.;
RT "Structural organization of the mouse metabotropic glutamate receptor
RT subtype 3 and its regulation by growth factors in cultured cortical
RT astrocytes.";
RL J. Biochem. 126:889-896(1999).
DR EMBL; AF170701; AAF06741.1; -.
DR EMBL; AF170697; AAF06741.1; JOINED.
DR EMBL; AF170698; AAF06741.1; JOINED.
DR EMBL; AF170699; AAF06741.1; JOINED.
DR EMBL; AF170700; AAF06741.1; JOINED.
```

```
DR EMBL: AF209923; AAF72873.1; -.
DR InterPro: IPR002956; Bridge_of_7less.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF00003; 7tm_3; 1.
DR PRINTS: PR01223; BRIDOF7LESS.
KW Receptor.
SQ SEQUENCE 345 AA; 38791 MW; 14B09E4C5E9F9B49 CRC64;

Query Match 22.4%; Score 521.5; DB 4; Length 345;
Best Local Similarity 33.2%; Pred. No. 1.6e-35;
Matches 125; Conservative 64; Mismatches 124; Indels 63; Gaps 11;

QY 39 YNLCDSRGAWGIVLEAVAGAGIVTTFVLTILVASLPEFVQDTKKRSLLGTQVFFLLGTL 98
DQ 12 YFLCDAGPGWIIIESLAIGIVVTILLALLAFLLMRKIQDCSQMNVLPQLLLLSVL 71

QY 99 GLFCLVACVVKPDFSTCASRRFLFGLVFAICFSLAAHVAFALNFKNKGPRGWIFT 158
DQ 72 GLFGLTFAFIQLNQTAPRVYFLFGLVFAICFSLAAHASNLVLRGRVSRFQWTFILF 131

QY 159 VALLLTVEVINTVEWLIITLVRSGGGPGNAGWAVASPCAIAANMDFVMA 218
DQ 126 WTTILCIAIGCSLLQIIATEYVTLITRG-----MMFVNMTPCQL-NVDEWVL 173

QY 213 LIYVMLLLGAFGLGAWPALCGRYKRWKKGHVFLVLTATSVAIWVWVMTYGNKO-HN 271
DQ 174 LVYVFLMALTFVSKATFCGPCENKQHGRLIFATVLSIIWVWISMLLRGNPQFOR 233

QY 272 SPTWDDPTALALANAARVFLFVPEVSQVTKSSPEQ-SYQGMVPTRGVGYETILKE 330
DQ 234 QPWDDPVCIALTNARVFLVLLYIVELCILYRSRQECFLOGNACPV--TAYQ----- 286

QY 331 QKGOSMEVNAKFS-----MDEPVAAKRPVSPYSYNGQLTSTVQPTMALMHKVP 383
DQ 287 ---HSFOVENQELSRSDGAEEDVA-----LTSYGTPIQPTVDPTQ----- 327

QY 384 EGAYDIILPRATANSQ 399
DQ 328 ----ECFIPQAKUSPQ 339

RESULT 9
Q9JIL6 PRELIMINARY; PRT; 300 AA.
AC Q9JIL6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE ORPHAN G-PROTEIN COUPLED RECEPTOR.
GN GPRC5D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21210966; PubMed=11311935;
RA Brauner-Osborne H., Jensen A.A., Sheppard P.O., Brodin B.,
RA Krogsaard-Larsen P., O'Hara P.;
RT "Cloning and characterization of a human orphan family C G-protein
RT coupled receptor GPRC5D(1).";
RL Biochim. Biophys. Acta 1518:237-248(2001).
DR EMBL: AF218809; AAF72880.1; -.
DR MGD; MGI:1935037; Gprc5d.
DR InterPro: IPR002956; Bridge_of_7less.
DR Pfam: PF00003; 7tm_3; 1.
DR PRINTS; PR01223; BRIDOF7LESS.
KW Receptor.
SQ SEQUENCE 300 AA; 34488 MW; D9423064260A4C59 CRC64;
```

```
Query Match 21.0%; Score 487.5; DB 11; Length 300;
Best Local Similarity 38.4%; Pred. No. 9.3e-33;
Matches 109; Conservative 44; Mismatches 112; Indels 19; Gaps 4;

QY 39 YNLCDSRGAWGIVLEAVAGAGIVTTFVLTILVASLPEFVQDTKKRSLLGTQVFFLLGTL 98
DQ 12 YFLCDFNEGPAWIVLESIAVIGIVVTILLALLAFLLMRKYQDCSQMNVLPQLFLLAVL 71

QY 99 GLFCLVACVVKPDFSTCASRRFLFGLVFAICFSLAAHVAFALNFKNKGPRGWIFT 158
DQ 72 GLFGLTFAFIQLNQTAPRVYFLFGLVFAICFSLAAHASNLVLRGRVSRFQWTFILF 131

QY 159 VALLLTVEVINTVEWLIITLVRSGGGPGNAGWAVASPCAIAANMDFVMA 218
DQ 132 IAGVSLQTIATIEYVTLITRG-----LMFEHMTPIQL-NVDEVCULLIVLF 179

QY 219 LLLGAFGLGAWPALCGRYKRWKKGHVFLVLTATSVAIWVWVMTYGNKO-HNSPTWDD 277
DQ 180 LMALTFVSKATFCGPCENKQHGRLIFATVLSIIWVWISMLLRGNPQORPHWDD 239

QY 278 PTLAIALAANAARVFLFVPEVSQVTKSSPEQSYQGMVPTRG 321
DQ 240 AVICIGLVTNAWVFLIYIPELSILYRSRQEC-----PTQG 277

RESULT 10
Q9BSPO PRELIMINARY; PRT; 105 AA.
AC Q9BSPO;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE SIMILAR TO G PROTEIN-COUPLED RECEPTOR, FAMILY C, GROUP 5, MEMBER
DE C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PANCREATIC ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC004925; AAH04925.1; -.
KW Receptor.
SQ SEQUENCE 105 AA; 11107 MW; D8B3FA9FA2559C99 CRC64;

Query Match 20.6%; Score 478; DB 4; Length 105;
Best Local Similarity 30.3%; Pred. No. 1.6e-32;
Matches 93; Conservative 2; Mismatches 4; Indels 4; Gaps 1;

QY 1 MAIHKALVMCLGLPLFLFPGAWAGHVPPGCSQGLNPLYYNLCDSRGAWGIVLEAVAG 60
DQ 1 MAIHKALVMCLGLPLFLFPGAWAGHVPPGCSQGLNPLYYNLCDSRGAWGIVLEAVAG 60

QY 61 IVTTFVLTILVASLPEFVQDTKKRSLLGTQV----FFLLGTGL 99
DQ 61 IVTTFVLTILVASLPEFVQDTKKRSLLGTQLRGRCHHTAGTMG 103

RESULT 11
Q9JMF0 PRELIMINARY; PRT; 262 AA.
AC Q9JMF0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE UNKNOWN PROTEIN (FRAGMENT).
GN GPRC5B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```



```

Db 118 RFLWGVLFAICFSCLLSQAMRVRRLRHGTPAGWQLVGLALCMLVQVIAVEMVLTV 177
QY 180 VRSGEGGPGNSSAGWAVASPCAIANMDFVMAIYVMLLLGALGAWPALCGRYKRW 239
Db 178 LRD-----RPACAYEPDMFVMAIYVMLLVVTLGLAFTLCGKFRWK 222
QY 240 KHGVFLLTATSVIAIWMVIMTYGN-KOHNSPTWDDPTLAIALAANAWAFVLFVVP 298
Db 223 LNAFLITAFSLVLIWAWMTYLFNGVKLQOGDANDPTLAIITLAASGWVFIHAIP 282
QY 299 EV-----SOVTKSSP---EQSYQGDMPYTRGVGYETILKEQKQSMFVENKA 342
Db 283 EIHTLLPALQENTPNYFDTSPMRMETAPEEDVQLPRA-----YMENKA 327
QY 343 FSDMEPVAA-----KRPVSPSYSGYNGQLLTSVYQPTMALM 378
Db 328 FSDMEHNAALRTAGFPNGSLGKRPSGLGRPSAPER-----SNVYQPTMAV 376

RESULT 4
ID 075205 PRELIMINARY; PRT; 403 AA.
AC 075205;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 44.8 KDA PROTEIN (GPCR5B PROTEIN).
GN A-69G12.1 OR GPCR5B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Adams M.D., Loftus B.J., Zhou L., Crosby M., Fuhrmann J., Mason T.M.,
RA Brandon R., Kim U.J., Kerlavage A.R., Venter J.C.;
RA "Homo sapiens Chromosome 16 BAC clone C1987SK-A-69G12.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Wei H., Osborne B., Spruyt M., Murphy D.;
RT "Cloning of a novel G protein-coupled receptor localized on human
RT chromosome 16p12.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Robbins M.J., Michalovich D., Hill J., Calver A.R., Medhurst A.D.,
RA Gloger I., Sims M.A., Middlemiss D.N., Pangalos M.N.;
RT "Molecular cloning and characterisation of two novel retinoic acid
RT inducible orphan G-protein coupled receptors.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004131; AAC27544.1; -.
DR EMBL; AF181862; AAF05331.1; -.
DR EMBL; AJ276101; CAC00632.1; -.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm3; 1.
DR PROSITE; P550259; G_PROTEIN_RECIP_F3_4; 1.
KW Hypothetical protein; Receptor.
SQ SEQUENCE 403 AA; 44795 MW; 3902A16C4F69C26E CRC64;

```

```

Query Match 31.5%; Score 733; DB 4; Length 403;
Best Local Similarity 39.4%; Pred. No. 4.4e-53;
Matches 164; Conservative 60; Mismatches 114; Indels 78; Gaps 9;

QY 1 MATHKALVCLGLFLFPG-AMAOGHVPPGCSQGLNPLYNLCDSGANGVILEAVAGA 59
Db 9 MRAHQVLTFL---LPIVTSVASENASTSRGCGLDLLPQVYSLCDLAIWGIIVEAVAGA 65
QY 60 GIYTFVLTILVASLPFVQDTKKRLLGTLGLFLGLFACVVKPDEFSTCASR 119

```

```

Db 66 GALITLLMLLLVRLPFIRKEKKKSPVGLHFLGLTGLGFLGTAFAIIQEDETICSVR 125
QY 120 RELFGVLFALCFSCLAHVFALNFLARKNHGPRGWIFTVALLTLTVEIINTEMLIITL 179
Db 126 RFLWGVLFALCFSCLLSQAMRVRRLRHGTPAGWQLVGLALCMLVQVIAVEMVLTV 185
QY 180 VRSGEGGPGNSSAGWAVASPCAIANMDFVMAIYVMLLLGALGAWPALCGRYKRW 239
Db 186 LRD-----RPACAYEPDMFVMAIYVMLLVVTLGLAFTLCGKFRWK 230
QY 240 KHGVFLLTATSVIAIWMVIMTYGN-KOHNSPTWDDPTLAIALAANAWAFVLFVVP 298
Db 231 LNAFLITAFSLVLIWAWMTYLFNGVKLQOGDANDPTLAIITLAASGWVFIHAIP 290
QY 299 EV-----SOVTKSSP---EQSYQGDMPYTRGVGYETILKEQKQSMFVENKA 342
Db 291 EIHTLLPALQENTPNYFDTSPMRMETAPEEDVQLPRA-----YMENKA 335
QY 343 FSDMEPVAA-----KRPVSPSYSGYNGQLLTSVYQPTMALM 378
Db 336 FSDMEHNAALRTAGFPNGSLGKRPSGLGRPSAPER-----SNVYQPTMAV 384

RESULT 5
ID 092320 PRELIMINARY; PRT; 410 AA.
AC 092320;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ORPHAN G PROTEIN-COUPLED RECEPTOR.
GN RAIG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Tao Q., Lotan R.;
RT "Molecular cloning and characterization of mouse retinoic acid-
RT inducible orphan G protein-coupled receptors.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF378831; AAK58076.1; -.
KW Receptor.
SQ SEQUENCE 410 AA; 45899 MW; E1DA5283270FF34 CRC64;

```

```

Query Match 30.8%; Score 716.5; DB 11; Length 410;
Best Local Similarity 40.7%; Pred. No. 1.1e-51;
Matches 162; Conservative 62; Mismatches 119; Indels 55; Gaps 9;

QY 13 LPFLFPGMAOQHVP--GCSQGLNPLYNLCDSGANGVILEAVAGAGIVTFTVLTII 70
Db 17 LPLLLVIASVASENASTSRGCGLDLLPQVYSLCDLAIWGIIVEAVAGALITLLMLJ 76
QY 71 LVASLPVQDTKKRLLGTQVFFLGTGLFLGLFACVVKPDEFSTCASRFLGVLPAIC 130
Db 77 LLVRLPPIKDKERRKRPVCLHFLGLTGLGFLGTAFAIIQEDETICSIRFLMGVLFALC 136
QY 131 FSLAAHVFALNFLARKNHGPRGWIFTVALLTLTVEIINTEMLIITLVRSGEQPG 190
Db 137 FSCLLSQAMRVRRLVRQGTSPASQVLVSLALCMLVQVIAVEMVLTVLRDT----- 189
QY 191 NSSAGWAVASPCAIANMDFVMAIYVMLLLGALGAWPALCGRYKRWKHFVLLTTA 250
Db 190 -----KPCAYEPDMFVMAIYVMLLVVTLGLAFTLCGKFRWKVNGAFILVTF 241
QY 251 TSAIWMVIMTYGN---KOHNSPTWDDPTLAIALAANAWAFVLFVYIPEVSQVTKSS 307
Db 242 LSAIWMVIMTYLFNGSLIKQGD---WSDPTLAIITLAASGWVFIHAIPFIHYTLTP 298
QY 308 PEOSYQGDMPY-----RGVGYETILKEQKQSMFVENKAFSMDPEVAAKRPVSPSY 362

```

Best Local Similarity 100.0%; Pred. No. 6e-186;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIHKALVMCLGLPLFLFPGAWAGHVPVPGCSQGLNPLYYNLCDSRGAWGIVLEAVAG 60
DB 1 MAIHKALVMCLGLPLFLFPGAWAGHVPVPGCSQGLNPLYYNLCDSRGAWGIVLEAVAG 60

QY 61 IVTFVLTIIIVASLPFVQDTKKRSLGTVQVFFLLGLFCLFVACVVKPDFSTCASRR 120
DB 61 IVTFVLTIIIVASLPFVQDTKKRSLGTVQVFFLLGLFCLFVACVVKPDFSTCASRR 120

QY 121 FLFGVLEAFCSCLAHVFALNFKRNHGRGWVITVALLLTLVEVIINTEWLIITLV 180
DB 121 FLFGVLEAFCSCLAHVFALNFKRNHGRGWVITVALLLTLVEVIINTEWLIITLV 180

QY 181 RGSCEGPGQNSSAGWAVASPCAIANDFVMAIYVMLLLGAFLGAWPALCGRYKRWK 240
DB 181 RGSCEGPGQNSSAGWAVASPCAIANDFVMAIYVMLLLGAFLGAWPALCGRYKRWK 240

QY 241 HGVEVLLTATSVAIWMYTYGNKQNSPTDPTTALAIALAANAWAFVLFYVPIEV 300
DB 241 HGVEVLLTATSVAIWMYTYGNKQNSPTDPTTALAIALAANAWAFVLFYVPIEV 300

QY 301 SQVTKSSPEQSYQDMYPTRGVGYETILKEQKQSMFVENKAFSMDPEPVAAKRPVSPYSG 360
DB 301 SQVTKSSPEQSYQDMYPTRGVGYETILKEQKQSMFVENKAFSMDPEPVAAKRPVSPYSG 360

QY 361 YNGOLLTSVYOPTEMALMHKVPSEGAYDIILPRATANSQVMSANSTLRAEDMYSAOSHQ 420
DB 361 YNGOLLTSVYOPTEMALMHKVPSEGAYDIILPRATANSQVMSANSTLRAEDMYSAOSHQ 420

QY 421 AATPPKDGKNSQVFRNPYVMD 441
DB 421 AATPPKDGKNSQVFRNPYVMD 441

RESULT 2

Q9N084 ID Q9N084 PRELIMINARY; PRT; 441 AA.
AC Q9N084;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE GPRC5C.
GN GPRC5C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Robbins M.J., Michalovich D., Hill J., Calver A.R., Medhurst A.D.,
RA Gloger I., Sims M.A., Middlemiss D.N., Pangalos M.N.;
RT "Molecular cloning and characterisation of two novel retinoic acid-
RT inducible orphan G-protein-coupled receptors (GPRC5B and GPRC5C).";
RL Genomics 76:8-18(2000).
DR EMBL; AJ276102; CAC00633.1; -.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR PROSITE; PS0259; G_PROTEIN_RECEP_F3_4; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 441 AA; 48179 MW; F541B4DED8437975 CRC64;

Query Match 100.0%; Score 2325; DB 4; Length 441;
Best Local Similarity 99.8%; Pred. No. 7.3e-186;
Matches 440; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIHKALVMCLGLPLFLFPGAWAGHVPVPGCSQGLNPLYYNLCDSRGAWGIVLEAVAG 60
DB 1 MAIHKALVMCLGLPLFLFPGAWAGHVPVPGCSQGLNPLYYNLCDSRGAWGIVLEAVAG 60

QY 61 IVTFVLTIIIVASLPFVQDTKKRSLGTVQVFFLLGLFCLFVACVVKPDFSTCASRR 120
DB 61 IVTFVLTIIIVASLPFVQDTKKRSLGTVQVFFLLGLFCLFVACVVKPDFSTCASRR 120

QY 121 FLFGVLEAFCSCLAHVFALNFKRNHGRGWVITVALLLTLVEVIINTEWLIITLV 180
DB 121 FLFGVLEAFCSCLAHVFALNFKRNHGRGWVITVALLLTLVEVIINTEWLIITLV 180

QY 181 RGSCEGPGQNSSAGWAVASPCAIANDFVMAIYVMLLLGAFLGAWPALCGRYKRWK 240
DB 181 RGSCEGPGQNSSAGWAVASPCAIANDFVMAIYVMLLLGAFLGAWPALCGRYKRWK 240

QY 241 HGVEVLLTATSVAIWMYTYGNKQNSPTDPTTALAIALAANAWAFVLFYVPIEV 300
DB 241 HGVEVLLTATSVAIWMYTYGNKQNSPTDPTTALAIALAANAWAFVLFYVPIEV 300

QY 301 SQVTKSSPEQSYQDMYPTRGVGYETILKEQKQSMFVENKAFSMDPEPVAAKRPVSPYSG 360
DB 301 SQVTKSSPEQSYQDMYPTRGVGYETILKEQKQSMFVENKAFSMDPEPVAAKRPVSPYSG 360

QY 361 YNGOLLTSVYOPTEMALMHKVPSEGAYDIILPRATANSQVMSANSTLRAEDMYSAOSHQ 420
DB 361 YNGOLLTSVYOPTEMALMHKVPSEGAYDIILPRATANSQVMSANSTLRAEDMYSAOSHQ 420

QY 421 AATPPKDGKNSQVFRNPYVMD 441
DB 421 AATPPKDGKNSQVFRNPYVMD 441

RESULT 3

Q9NZHO ID Q9NZHO PRELIMINARY; PRT; 395 AA.
AC Q9NZHO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ORPHAN G-PROTEIN COUPLED RECEPTOR.
GN GPRC5B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Brauner-Osborne H., Krogsgaard-Larsen P.;
RT "Sequence and expression pattern of a novel human orphan G-protein-
RT coupled receptor, GPRC5B, a family C receptor with a short amino-
RT terminal domain.";
RL Genomics 65:121-128(2000).
DR EMBL; AF202640; AAF67321.1; -.
DR InterPro; IPR002956; Bridge_of_7less.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR PRINTS; PR01223; BRIDGOF7LESS.
DR PROSITE; PS0259; G_PROTEIN_RECEP_F3_4; 1.
KW Receptor.
SQ SEQUENCE 395 AA; 43845 MW; 40ACD233038811D4 CRC64;

Query Match 31.5%; Score 733; DB 4; Length 395;
Best Local Similarity 39.4%; Pred. No. 4.3e-53;
Matches 164; Conservative 60; Mismatches 114; Indels 78; Gaps 9;

QY 1 MAIHKALVMCLGLPLFLFPG-AWAAGHVPVPGCSQGLNPLYYNLCDSRGAWGIVLEAVAG 59
DB 1 MRAHQVITFLL- --LFVITSVASENASTSRGCGLDLLPQVYVSLCDLDAIMGIVVEAVAG 57

QY 60 GIVTFVLTIIIVASLPFVQDTKKRSLGTVQVFFLLGLFCLFVACVVKPDFSTCASRR 119
DB 58 GALTITLLMLLLVRLPFIFKEKEKSPVGLHFLGLTGLTGLTFAFIQEDETICSVR 117

QY 120 RFLFGVLEAFCSCLAHVFALNFKRNHGRGWVITVALLLTLVEVIINTEWLIITL 179

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2002, 09:32:51 ; Search time 31.28 seconds
(without alignments)
2438.963 Million cell updates/sec

Title: US-09-895-686-1
Perfect score: 2326
Sequence: 1 MAHKALVMCLGLPLFLPG.....ATPPKDGKNSQVFRNPYWD 441

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeop.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2326	100.0	441	4 Q9NZG5	Q9nzg5 homo sapien
2	2325	100.0	441	4 Q9NQ84	Q9nq84 homo sapien
3	733	31.5	395	4 Q9NZH0	Q9nzh0 homo sapien
4	733	31.5	403	4 Q75205	Q75205 homo sapien
5	716.5	30.8	410	11 Q92320	Q92320 mus musculus
6	651	28.0	162	4 Q9NXX0	Q9nxx0 homo sapien
7	558	24.0	357	4 Q95357	Q95357 homo sapien
8	521.5	22.4	345	4 Q9NZD1	Q9nzd1 homo sapien
9	487.5	21.0	300	11 Q9JIL6	Q9jil6 mus musculus
10	478	20.6	105	4 Q9BSP0	Q9bsp0 homo sapien
11	425.5	18.3	262	11 Q9JMF0	Q9jmf0 mus musculus
12	167	7.2	1404	5 Q20073	Q20073 caenorhabditis
13	160.5	6.9	879	11 Q9QYS2	Q9qys2 mus musculus
14	155	6.7	917	5 Q24265	Q24265 drosophila
15	154	6.6	738	5 Q9V4U3	Q9v4u3 drosophila
16	151	6.5	872	4 Q9H3N6	Q9h3n6 homo sapien

17	150.5	6.5	877	4 Q9UGT0	Q9ugt0 homo sapien
18	143	6.1	1199	11 Q9EPV6	Q9epv6 mus musculus
19	140	6.0	589	4 Q9UGS9	Q9ugs9 homo sapien
20	133	5.7	1218	13 Q90ZF3	Q90zf3 oncorhynchus
21	132.5	5.7	551	13 Q9PWQ0	Q9pwq0 fugu rubrip
22	132	5.7	251	13 Q73657	Q73657 fugu rubrip
23	129.5	5.6	251	13 Q73656	Q73656 fugu rubrip
24	129	5.5	583	5 Q9VNZ5	Q9vnz5 drosophila
25	128.5	5.5	877	13 Q9PW88	Q9pw88 carassius a
26	123	5.3	875	13 Q73640	Q73640 fugu rubrip
27	121.5	5.2	251	13 Q73653	Q73653 fugu rubrip
28	119.5	5.1	251	13 Q73654	Q73654 fugu rubrip
29	119.5	5.1	251	13 Q73655	Q73655 fugu rubrip
30	119	5.1	250	13 Q73646	Q73646 fugu rubrip
31	117.5	5.1	699	16 Q9ABL0	Q9abl0 caulobacter
32	117	5.0	864	13 Q73637	Q73637 fugu rubrip
33	116	5.0	940	13 Q73635	Q73635 fugu rubrip
34	115.5	5.0	2426	3 Q96UR0	Q96ur0 aspergillus
35	115	4.9	458	13 Q93555	Q93555 carassius a
36	115	4.9	856	13 Q73638	Q73638 fugu rubrip
37	115	4.9	983	11 Q62916	Q62916 rattus norv
38	114	4.9	844	13 Q93552	Q93552 carassius a
39	113	4.9	250	13 Q73647	Q73647 fugu rubrip
40	113	4.9	848	13 Q93553	Q93553 carassius a
41	112.5	4.8	362	13 Q9PSY1	Q9psy1 carassius a
42	112.5	4.8	408	13 Q93558	Q93558 carassius a
43	112	4.8	390	2 Q9L949	Q9l949 pseudomonas
44	110.5	4.8	250	13 Q73650	Q73650 fugu rubrip
45	110	4.7	250	13 Q73651	Q73651 fugu rubrip

ALIGNMENTS

RESULT	1
Q9NZG5	
ID	Q9NZG5 PRELIMINARY; PRT; 441 AA.
AC	Q9NZG5;
DT	01-OCT-2000 (TREMBLrel. 15, Created)
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	ORPHAN G-PROTEIN COUPLED RECEPTOR (G PROTEIN-COUPLED RECEPTOR, FAMILY
DE	C, GROUP 5, MEMBER C).
GN	GPC5C.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21210966; PubMed=11311935;
RA	Brauner-Osborne H., Jensen A.A., Sheppard P.O., Brodin B.,
RA	Krogsgaard-Larsen P., O'Hara P.;
RT	"Cloning and characterization of a human orphan family C G-protein
RT	coupled receptor GPRC5D(1).";
RL	Biochim. Biophys. Acta 1518:237-248(2001).
RN	[2]
RC	SEQUENCE FROM N.A.
RC	TISSUE=PANCREAS, EPITHELIOID CARCINOMA;
RA	Strausberg R.;
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF207989; AAF72870.1; -.
DR	EMBL; BC016860; AAH16860.1; -.
DR	InterPro; IPR003337; GPCR_Mgr.
DR	InterPro; IPR003006; Ig_MHC.
DR	Pfam; PF00003; 7tm3.1.
DR	PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
DR	PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW	Receptor.
SQ	SEQUENCE 441 AA; 48193 MW; B255D9680328FE3D CRC64;
Query Match	100.0%; Score 2326; DB 4; Length 441;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 14:34:17 : Search time 1947.48 Seconds
(without alignments)
12606.528 Million cell updates/sec

Title: US-09-895-686-7

Perfect score: 1819
Sequence: 1 cgcctcgcgcctccaccacgc.....ctttatctactcttaaaaa 1819

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 674847542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gD_est1:*
10: gD_est2:*
11: gD_hic:*
12: gD_gss:*
13: em_gss_hum:*
14: em_gss_iny:*
15: em_gss_pla:*
16: em_gss_vtl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1081	59.4	1815	11 AK014308	AK014308 Mus muscu
2	885.6	48.7	903	9 AL545009	AL545009 AL545009
3	882.8	48.5	955	9 AL570913	AL570913 AL570913
4	852.6	46.9	906	9 AL577178	AL577178 AL577178
5	847	46.6	874	9 AL556748	AL556748 AL556748
6	845.2	46.5	905	10 BI916185	BI916185 603178072
7	837.6	46.0	851	9 AL546406	AL546406 AL546406
8	792.6	43.6	842	9 AL571754	AL571754 AL571754
9	775.2	42.6	956	10 BG829477	BG829477 602763774
10	761.2	41.8	846	9 AL578937	AL578937 AL578937
11	747.8	41.1	873	10 BI769602	BI769602 603054983
12	744.4	40.9	820	10 BI834606	BI834606 603089872
13	732	40.2	933	10 BI772730	BI772730 603053260
14	728.4	40.0	776	10 BG750651	BG750651 602708511
15	710.2	39.0	1052	10 BI822698	BI822698 603036033
16	687	37.8	689	10 BI764504	BI764504 603050790
17	678.8	37.3	710	9 AM954136	AM954136 EST366206

18	607.2	33.4	685	10 BE898354	BE898354 601681172
19	588.8	32.4	623	9 AA877534	AA877534 nc01608.s
20	558.4	30.7	698	10 BI760442	BI760442 603045457
21	529	29.1	653	10 BI400350	BI400350 MI-P-AV1-
22	517.2	28.4	849	10 BE909608	BE909608 601502794
23	479.2	26.3	580	10 BI404524	BI404524 MI-P-CP1-
24	478.6	26.3	575	10 BE683495	BE683495 183298 MA
25	475	26.1	478	9 AM959466	AM959466 EST371536
26	472.4	26.0	588	10 BG692011	BG692011 341537 BA
27	459	25.2	960	9 AI921413	AI921413 w024e08.x
28	458.2	25.2	900	10 BF786181	BF786181 602110463
29	455.6	25.1	959	10 BF783102	BF783102 602109312
30	455.8	25.1	459	9 AI075186	AI075186 oy96b09.x
31	451.6	24.8	555	10 BI402962	BI402962 MI-P-CP1-
32	443	24.4	443	9 AI298839	AI298839 qm93h08.x
33	441.6	24.3	548	9 AA178961	AA178961 zp59f04.s
34	436	24.0	436	9 AI659014	AI659014 tu23f06.x
35	431.8	23.7	443	9 AI016010	AI016010 ou95h01.x
36	427.2	23.5	444	9 AI694173	AI694173 wc83f12.x
37	427.2	23.5	513	10 BM107227	BM107227 511394 MA
38	423.2	23.3	566	9 BE232134	BE232134 137070 MA
39	422.4	23.2	432	9 AI052635	AI052635 oz18a01.x
40	422	23.0	510	10 BI403847	BI403847 MI-P-CP1-
41	419.2	23.0	550	10 BG082452	BG082452 H3076D03-
42	416.6	22.9	522	10 BI338854	BI338854 363561 MA
43	415	22.8	503	10 BG610066	BG610066 324540 MA
44	408.6	22.5	606	10 BI825005	BI825005 603035547
45	405.4	22.3	442	9 AA902120	AA902120 ok94e04.s

ALIGNMENTS

RESULT 1	AK014308	1815 bp	mus musculus 14, 17 days embryo head cDNA, enriched library, clone:320002M13:homolog to ORFHAN G-PROTEIN
LOCUS	AK014308		
DEFINITION	AK014308	1815 bp	mus musculus 14, 17 days embryo head cDNA, enriched library, clone:320002M13:homolog to ORFHAN G-PROTEIN
ACCESSION	AK014308	GI:12852069	
VERSION	AK014308.1		
KEYWORDS	HTC: CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE	3 (sites)		
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitsuani,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujikake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Mataliki,M., Onoda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunaga,S., Kawai,D., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		

MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

20530913
 11076861
 4 (sites)
 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 5 (bases 1 to 1815)
 Adachi, J., Alizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
 Hume, D., Imoto, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Kono, H., Konda, M., Koya, S.,
 Kuhlberg, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
 Nunazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Ouackenbush, J.,
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D.,
 Schiraldi, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,
 Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
 Hayashizaki, Y.

TITLE
JOURNAL

Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki. The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

COMMENT

Please visit our web site (http://genome.gsc.riken.go.jp/) for
 further details.

cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues. First strand cDNA was primed with a primer
 [5' GAGAGAGAGAGATCCAGACGCTCTTTTCTTTTCTTNN 3'], cDNA was
 prepared by using trehalose thermo-activated reverse transcriptase
 and subsequently enriched for full-length by cap-trapper. Second
 strand cDNA was prepared with the primer adapter of sequence [5'
 GAGAGAGATCTTCAGTAAATTAATTAATTCCTCCCTCC 3']. cDNA was cleaved
 with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
 Host: SOLR.

FEATURES
source

Location/Qualifiers
 1..1815
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="MGD:102714"
 /db_xref="taxon:10090"
 /clone="3200002M13"
 /tissue_type="head"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dex_stage="14, 17 days embryo"
 1..1815
 /note="data source:SPRR, source key:09NZG5, evidence:ISS
 homolog to ORPHAN G-PROTEIN COUPLED RECEPTOR"
BASE COUNT
 342 a 544 c 503 g 426 t
ORIGIN

Query Match 59.4% Score 1081; DB 11; Length 1815;
 Best Local Similarity 79.4%; Pred. No. 1,1e-218;
 Matches 1415; Conservative 0; Mismatches 335; Indels 33; Gaps 10;

Qy 48 aggaagccaacagagctgctgagagccagagatgcatcacaagcctggtat 107
 |||||
 Db 42 AGGACCAAGCTGAAGCTGGCTGGAGCCAGATGGCCACTCACAACAACCTGAT 101
 |||||
 Qy 108 gtgcctgagctgctctctctcctgcttcacagggcctgagccagagcctatgccacc 167
 |||||
 Db 102 GTGCCTAGAGCTGCTCTCTTC---TTCCAGAGAGCCTTGCCCAAGATCATGCCCAACC 158
 |||||

Qy 168 cggctgcagccaagagcctcaacccctctactacaacctgtgtaccgctctggggcgtg 227
 |||||
 Db 159 TGGTTGACAGCCAGATCTGGATCCCTCTACTACAACCTTGATGACCGCTGGGGGCTG 218
 |||||
 Qy 228 gggcagctgctctgagagcgcgtgtgctggggcgagcatgtgtacacagttgtgtccacat 287
 |||||
 Db 219 GGGCATTTCTTGGAGGAGGAGTGGGAGGAGGATCATCAGACATTTGTGTAAACCAT 278
 |||||
 Qy 288 catctgtgtgacagcccccctctctgtgcaagacccaagaaagagcctggggagac 347
 |||||
 Db 279 CATCTTTGTGGCTAGACCTTCCATTTTGTGACAGACACTAAGAGCGGA-CTCTTGGGGAC 337
 |||||
 Qy 348 ccaggtatctctctctctctgaggaacccctgagccctctctgctgtgtgtgtgtgtgt 407
 |||||
 Db 338 CCAAGTGTCTTCCCTGCTGGGACACCTGGGCTGCTCTTGTGCTCGATTTGGCTGTGGT 397
 |||||
 Qy 408 gaagccagcatctctccacctgtgctctctcggcgctctctctgttggttcgttcgcgat 467
 |||||
 Db 398 GAAGCCGAGCTTCTTACCTGTGCTGCTGCTGACGCTTCTCTTGGGGTCTGTTGGCAT 457
 |||||
 Qy 468 ctgctctctctgtctgagcgctcaagctctctcctcaactctctgcccggagaaacca 527
 |||||
 Db 458 CTGCTTCTCTGCTGTGGTGAAGCTCAAGCTCTTCTCTCACTTCTTAACCCGGAAGAACCA 517
 |||||
 Qy 528 cgggcccggggcctggtgtgacttcaactgtgctctgctgtagaccctgtagagtgat 587
 |||||
 Db 518 TGGGCCCCGAGAGCTGGGATCTTACCGGTGGCGCTGCTGCTCAACCTTGGAGGTCAT 577
 |||||
 Qy 588 catcaatcagaagtgctgtgatacaaccctgtgtcggggagctgagggcgagccctca 647
 |||||
 Db 578 CATTAACACCGAGTGGCTTATCATCATCCCTGTACGGGGAGTGGGACGTTTACCCCT 637
 |||||
 Qy 648 gggcaacagcagcagcagcgtggcgctggcgctcccccctgcatcagcgaactgagact 707
 |||||
 Db 638 GGGCAATTCAGTGCAGCTGCAGCTGCAGCATACCTCTCCGCTGCATCCGCAACATGAGACTT 697
 |||||
 Qy 708 tgcatagtcactcaactcaagctgctgtgctgtgctgagggctctcctggggcctgtg 767
 |||||
 Db 698 TGTATGCTCTCTCATCTACGTAATGCTGTTGCTCTGACGGCTTCTTACGAGAGCTGGCC 757
 |||||
 Qy 768 ggcctgtgtgagcgcctcaagagcgtggagaaagatgaggtctctgtgtcccaacac 827
 |||||
 Db 758 CACCTGTGTGGCGGCTTCAAGCGCTGGGGAACAGCGGGGTCTTGTGTGCTCACAC 817
 |||||
 Qy 828 agcaacctcgtgtgcatatggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 887
 |||||
 Db 818 TGTATCTCCATTCGCATCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 877
 |||||
 Qy 888 gcaacaacagctccacacgtggatgtagaccacacgtgtgcatcgtcgcgcgaatgctg 947
 |||||
 Db 878 GCACCATATGCCCCACCTGGGATGACCCACCTTGGCCATTCCTCGCTGCATATGCTCG 937
 |||||
 Qy 948 ggcctgtgctctctctcaagctacccagagctgtccaggtgagcgaagttccaaaccaga 1007
 |||||
 Db 938 GACCTTGTCTCTTCTATGTGTAT-CTGTAGGCTTCAAGGTGACAAACTTACCCAGA 996
 |||||
 Qy 1008 gcaagctacccaagggagcatgttaccaccaaccggggcgctggtgcatgtagaccatcgtgaa 1067
 |||||
 Db 997 ACAGAGCTACCA-GGGGACATGTACCCGACCCGAGGGGTGGCTATAGACATCTCGAA 1055
 |||||
 Qy 1068 agagcagaaggtlcaagagcatgtctgtgagaacaagccttctcatgtgagtcagcgt 1127
 |||||
 Db 1056 GGACAGCAGCGGGCCAGAGCATGTTGTGGAGAACAGGATTTCTATGTGACAAACAGC 1115
 |||||
 Qy 1128 tgcagctcaagggcggtgtgtaccatcaagcggttacaatgaggtgagctgtcgtgaccagt 1187
 |||||
 Db 1116 CTCAGCAAGAGACACAGTGTACCTTACAGTGGCTTAAGTGAAGCTGTGACACACGT 1175
 |||||
 Qy 1188 gtaccagcccatgagatgtgcccgtgtagcaaaagcttcgcgcgaagagatcttcagcat 1247
 |||||
 Db 1176 GTACCAAGCCACGACAGATGGCTGTGATGACAAAGGCCGCTCTAAGAGTGTGATACGAGT 1235
 |||||
 Qy 1248 catctctccacagggcgaacggcgaacagcaggtgtgtgtgagcagtgccaaatgcagcctgcg 1307
 |||||

Db	1236	CATCCTCCACGAGGCGCACCCGCCAACAGCCAGGTGATGGGGAGTGGCAACTCAACCTTCGCG	1295
Qy	1308	ggctgaagaacatgtaactcgcgcccaagaaccacacagcgagccacacccgcgaaagacgagca	1367
Db	1296	GGCCGAAACATATGTAACATAGTGTCACAGAGCCACACAGTGGCCACGACCAACAAAGACGCGAA	1355
Qy	1368	gaactctgaagctctttaaaaccctcagctgtgagacgaagtcagc-----gttgcgag	1423
Db	1356	GATCTCTCAGGCTCTTTAAAAATCCCTACGTGTGGAGCTAAAGGACACAGCTGTGGCCAGGA	1415
Qy	1424	agagcgcgctcgatttgggagagcgccctgaaggaactcggcccgagcaagagactccag	1483
Db	1416	GGAAGCAGGTTGATTGTACAGAGGGCCCTTGGGACCTGGTTGGTGAGGAGACACACAGG	1475
Qy	1484	gctcct-----cctccccctcgagagccacgaacatgycgccagaatg	1530
Db	1476	AGCCCATGCTCTATNGGCTCCCTCCCTCTTACTACTGTTCAAGGCTAGCTTATGCTCAAGG	1535
Qy	1531	aagggccctccctctcgtgcagatgttgggtgggtgtaatggtgtgccccacccactctc	1590
Db	1536	TCAGAGCCCTCGATCTGCTGATGTTTGGTGGGTGTTCATGAGGGCCCTCCCA--CCG	1592
Qy	1591	agctgttggtagtcgaagagcaaccacccacgctcctctgcagatcacactcggcgctac	1650
Db	1593	AGTATTGTGGAGTCCACAGGCCCTGCGCTGTGTTCAATGCGAGGCCACCTTGCTATCTA	1652
Qy	1651	actccagcaaatagtgcttc--ggggtgtgtgcctgagcgactatgctctcgtg	1707
Db	1653	ACTCATATCAATATAGCGTTCTGCTGGGATGGTGCTGTCTATGCTACGCTCTCCGG	1712
Qy	1708	agattccgcga--cctcaagagactcccaagcgctagc---gctgagatctgtctc	1763
Db	1713	TGACACCTATAGGCTCCATTGAGCTCAGAGAGGATGCTTAAGGCTGAGTCTGCTCTC	1772
Qy	1764	tgtgagacaagggctcctaataataatcatcttcgcttalt	1806
Db	1773	TGTGAGGAGTAAAGGTGCTTATAATATCATTTCTGCTTAT	1815
RESULT	2		
LOCUS	AL545009	903 bp	mrna linear EST 16-FEB-2001
DEFINITION	AL545009 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D101YLL0 5		
ACCESSION	AL545009		prime, mRNA sequence.
VERSION	AL545009.1	GI:12877490	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 903)		
AUTHORS	Ll.W.B., Gruber,C., Jesse,J. and Polayes,D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 Evry cedex - France		
	Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
FEATURES	Location/Qualifiers		
source	1..903		

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1013Y110"
/clone_11b="LTI.NF1006.PL2"
/tissue_type="placenta"
/note="Vector: pcwmsport 6. Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcwmsport 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Jiang Life Technologies."

```

a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8377
Email : filang@lifetech.com URL :
<http://fulllength.invitrogen.com>"

Query Match	48.78;	Score 885.6;	DB 9;	Length 903;
Best Local Similarity	99.68;	Pred. No. 2.2e-177;		
Matches 888; Conservative	0;	Mismatches 4;	Indels 0;	Gaps 0;

[illegible]


```

RESULT      3
AL570913/c  955 bp  mRNA  linear  EST 16-FEB-2001
LOCUS       AL570913 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1013YL10 3
DEFINITION  prime. mRNA sequence.
ACCESSION   AL570913
VERSION     AL570913.1  GI:12927686
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 955)
AUTHORS     Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..955
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1013YL10"
/clone_id="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI. 1st strand cDNA
was primed with a NotI-oligo(dt) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

```

BASE COUNT  175 a 275 c 301 g 200 t 4 others
ORIGIN

```

```

Query Match      48.5%; Score 882.8; DB 9; Length 955;
Best Local Similarity 98.7%; Pred. No. 8.8e-177;
Matches 918; Conservative 3; Mismatches 5; Indels 4; Gaps 3;

QY 806 gggtcttgcctcctccacacagccctccgttcgcatatggtggtgagatcgca 865
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 936 GGGTCTTTGTCT-CTTCACACAGCCACCTCGGTGCGCATATGGGTGTGGATCGTCA 878
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 866 tctatcttaagcacacaagacagacaaagtcaccactggatgacccacagctggca 925
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 877 TGTATACTTACGGCAACAGCAGCAACAGTCCACCTGGGATGACCCGACGCTGGCA 818
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 926 tgcgcctgcgcgcaagtgcctgggcttcgtcctcttcttaagtcac -ccccgaagctcc 984
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 817 TCGGCTTCGCGCCCATCTCGGGCTTCGTCTTCTTACGTCAATCCCGGAGGCTTCC 758
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 985 caggctgaccaggtccagccagagaaagctaccacagggagacatctaccaccccggggc 1044
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 757 CAGGTGACCAAGTCCAGCCCGAGAGCAAGCTAACAGGGGAGACATGTACCCACCCGGGGC 698
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1045 gttgggtctatgagaccatctctaaagagagaaggttcagagacatgtctgtgagaacaag 1104
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 697 GTGGGCTATGAGACCATCTCGAAGAGCAGAAAGGTGTCAGAGCATGTTCTGTGAGAACAG 638
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1105 gcccttccatgagatgagcgggttcgagactaagagccggtgtcaccatacagcggttac 1164
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 637 GCGTTTTCATGAGATGACCGCGTTTCACCTTAAGAGCGCGGTCTCACCATACAGCGCGGTAC 578
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1165 aatggagcagctgctgaccagtggtgtaccagccacttgatgagcctgatacacaaatt 1224
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 577 AATGGGACAGCTGCTGACCAAGTGTGTACCAAGCCCACTGAGATGGCCCTGTATGACAAAGTT 518
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

QY 1225 ccgtccgaagagcttaccacalcatcctcccaaggccacccgcaacagctgtag 1284
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 517 CCGTCGAAGAGGCTTACGACATCATCTCCACAGGGCCACCGCCACAGCGAGGTAGT 458
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1285 ggcagtgccaacttgacccttgaggctgaagacatgtacttggtcccaagaccagagc 1344
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 457 GCGAGTGGCAATCTGA -CTGCGGGCTGAAAGACATGTACTCGGCCAGAGCCACAGGGC 400
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1345 gccacacccgcgaagaagacagcaaaccttcagctctttagaacccctagtgtagac 1404
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 399 GCCACACCGCCGCAAGAGAGCGCAAGAACTCTCAGGCTTTAGAAACCCCTACGTGTGGAC 340
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1405 tgaagtcagcgtgtgcgaggaagagcggtcgaatttgaggagggccctagagaccctgccc 1464
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 339 TGAGTCACGCGGTGCGAGAGAGAGCGCGTCCGATTTGGGAGAGGCCCTGAGCACTGGCCC 280
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1465 cgggcaagggactctccagagctccctccctcccttgagcagggcccaacaatgtgcccag 1524
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 279 CGGGCAAGGGACTCTCAGGCTCTCTCCCTGGCAGGCCACACATGTGCCAG 220
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1525 atgtgaagggcctccctctctgcagtgatttggtgagtgatcagtggtgtccaccaca 1584
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 219 ATGTGGAAGGCGCTCCTCTCTGCGCAGTGTGGGTGGTGTATGAGGTGTCACACCA 160
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1585 ctccctcagtgattgtgagagtcgagagagcaacccagcctctctgcagatcacctcgyc 1644
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 159 CTCTCAGTGTGTGTGAGTGCAGAGARCCAAACCACTCTTACAGAGAMCACTCGGC 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1645 ggtcacactccagccaatatgtctctcgagggtggtgctgagcagcgctatgtttctc 1704
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 99 GGTCACTCTCAACCAATAATGTTCTCGGGGTGGTGGCTGGGAGCCGCTATGTTCTTC 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 1705 tggagattcctgcaaccccaagagacttc 1734
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 39 TGGAGATTCTCTCAACCTCAAGATGATMCC 10
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT      4
AL577178/c  906 bp  mRNA  linear  EST 16-FEB-2001
LOCUS       AL577178 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1084YM03 3
DEFINITION  prime. mRNA sequence.
ACCESSION   AL577178
VERSION     AL577178.1  GI:12940052
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 906)
AUTHORS     Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..906
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1084YM03"
/clone_id="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site 1: NotI. 1st strand cDNA
was primed with a NotI-oligo(dt) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :

```


BASE COUNT 157 a 268 c 290 g 189 t 2 others
 ORIGIN

Query Match 46.9%; Score 852.6; DB 9; Length 906;
 Best Local Similarity 99.7%; Pred. No. 2.2e-170;
 Matches 874; Conservative 1; Mismatches 0; Indels 2; Gaps 2;

QY 871 acttaccgcaagaagcaagcaacagtcacacgtgagatgacccacgctgacatgcc 930
 |||||||
 DB 906 ACTTACGGCAACAGCAGCACAACAGTCCACCTGGGATACCCACGCTGGCATGCC 847
 QY 931 ctgcgcgaacatgctgagctctctctctctctctctctctctctctctctctct 990
 |||||||
 DB 846 CTCGCCGCATATCCTGGGCC-TCGTCTCTTCTTACCTCATCCCGAGGTCTCCAGAGT 788
 QY 991 accaagtcacagcccaagcaaaagctacacaggggagacatgacccacccgggctgtgac 1050
 |||||||
 DB 787 ACCAAGTCCAGCCCGCAGAGCAAAAGCTACACAGGGGACATGATCCCGCGGCGTGGC 728
 QY 1051 tatgagacacatctctgaaagagaaaggtcagaagcatgtctgtgagaaacagcctt 1110
 |||||||
 DB 727 TATGAGACCATCTCTGAAGAGCAAGAGGTCAAGAGCATGTTGCTGGAACAAGGCTTT 668
 QY 1111 tcatgtatgagccggtttgagctaaagagccggtgtcacacatagacgggtgacatggg 1170
 |||||||
 DB 667 TCCATGATGAGCGCGGTGAGCTAAGAGGCGCGGTGCACCATACACGCGGTCAATGGG 608
 QY 1171 cagctgtgacagctgtgtacacagccacatgagatgagcctgtatgcaaaagttccgtcc 1230
 |||||||
 DB 607 CACTGTGTGACAGTGTGTACAGCCACAGATGAGCTGATGATGCAAAAGTTCCGTCC 548
 QY 1231 gaagagcttaagacatcatctctcgaaggagccagccagccagcagggtgattggagct 1290
 |||||||
 DB 547 GAAGGAGCTTACACATCATCTCCACGCGGCAACCGCAGCGATGATGGAGT 488
 QY 1291 gcaaacctcagacctctgaggggtgaaagatatctcggagcccaagcccaagggcgccaca 1350
 |||||||
 DB 487 GCCAATCTCGA-CCCTGCGGGTGAAGACATGTACTGCGCCAGAGCCACAGCGCGCCACA 429
 QY 1351 ccgcgcgaagaagcaagaactctcaggtctctctctctctctctctctctctctctctctc 1410
 |||||||
 DB 428 CGCCCGAAGACGGCAAGACTCTCAGGTCTTTAAGAAACCTTACGTGTGGAGTGAAGTC 369
 QY 1411 aagcgttggcgaagagagagcggtcgtgatttgggagggccctgagagacctggccccggca 1470
 |||||||
 DB 368 AGCGGTGGCGAGAGAGAGGCGGTGGATTGGGAGAGGCGCTGAGAGACCTGGCCCGGCA 309
 QY 1471 aaggagctcagaggtct 1530
 |||||||
 DB 308 AGGAGACTCTCAGAGCTCTCTCTCCCTGCGAGGCCAGCAACATGTGCCCCAGATGTG 249
 QY 1531 aaggagctcctc 1590
 |||||||
 DB 248 AAGGGCTCTCCCTCTCTCTCAGTGTGGTGGGTGATGATGGGTGCCCAACCTCTCTC 189
 QY 1591 agtgttctgtgagatcgagagagccaaacccagcctctctgcaagatcaactctgagctgac 1650
 |||||||
 DB 188 AGGTGTGTGAGTTCGAGGAGCCAAACCCAGCTCTCTGCAAGATCACTCTGGGGGTAC 129
 QY 1651 acccagcgaacaaatagttctctgaggggtgtgtgagcagcgccctatgttctctctgaga 1710
 |||||||
 DB 128 ACTTCAGCCAAATATATGTCTTCGGGGTGTGGCTGGGACAGCCCTATGTTTCTTGAGAG 69
 QY 1711 ttcctgcaacctcaagaagacttcccaagcgctcagcg 1747
 |||||||
 DB 68 TTCCTGCAACCTCAAGAGACTTCCAGGCGCTAGGC 32

RESULT 5
 AL556748 AL556748 874 bp mRNA linear EST 16-FEB-2001
 LOCUS

DEFINITION AL556748 LTR_NFL006_PL2 Homo sapiens cDNA clone CS0DK005YN05 5
 prime, mRNA sequence.
 ACCESSION AL556748
 VERSION AL556748.1 GI:12899712
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 874)
 AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: segreff@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 FEATURES
 source
 1..874
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DK005YN05"
 /clone_1b="LTR_NFL006_PL2"
 /issue_type="placenta"
 /note="Vector: pCMVSPORT 6; Site: 1; NotI: 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,
 a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetec.com URL :
 http://fulllength.invitrogen.com"

BASE COUNT 128 a 281 c 264 g 200 t 1 others
 ORIGIN

Query Match 46.6%; Score 847; DB 9; Length 874;
 Best Local Similarity 99.3%; Pred. No. 3.3e-169;
 Matches 861; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 19 gccggaagctacagctgctcagctcagctgagggagcccaagcagcctgagctggagcc 78
 |||||||
 DB 1 GCCGGAAGTACGAGTGGCTCAGCTGAGAGGCCAACCAGAGCTGAGTGGAGCC 60
 QY 79 agatgagcatcccaaaagccttgatgagctgagctgagctcctctctctctccca 138
 |||||||
 DB 61 AGATGGCCATCCACAAGCCTTGATGTGCTGGAGTGCCTCTCTCTGTCCTCA 120
 QY 139 ggggctgggcccagggcattccacccggctgacagcccaaggctccaacccctgtac 198
 |||||||
 DB 121 GGGGCTGGGGCCAGGGCCATGTCCACCGCGTGCAGCCCAAGGCTCAACCCCTGTAC 180
 QY 199 tacaacctgtgacagctctcggggcggtggagcatctcctctggagggcggtggggcg 258
 |||||||
 DB 181 TACAACCTGTGTGACCCGCTGTGGGGCGTGGGGCATGTGCTGTGGAGCGGTGGGGCG 240
 QY 259 ggcattgtcaacaagttgtgtacacatcatctctgtgtgagcagcctcctctgtgag 318
 |||||||
 DB 241 GGCATGTGACACAGTTGTGCTCACCATCATCTGAGGCGCACGCTCTTGTGTGAG 300
 QY 319 gaccccaagaaagagcctgtctgggagccaggtatctctctctctctgggagccctggcg 378
 |||||||
 DB 301 GACACCAAGAAACGAGCCTGCTGGGACCCAGGATTTCTTCTGAGGACCTTGCGC 360
 QY 379 cctctctgacctgt 438
 |||||||
 DB 361 CTCTTCTGCGCTGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
 QY 439 cgcttctcttgggggtctgtgtgcacatctgtctctctgtgtgtgtgtgtgtgtgtgtgt 498
 |||||||
 DB 421 CGCTTCTCTTGGGGT 480

Y	499	gcccacacctccctcggcccggaagaacacacggcccgccggggtgggtgatacttcaactgtg	558
Y	499	gcccacacctccctcggcccggaagaacacacggcccgccggggtgggtgatacttcaactgtg	558
Db	481	GCCCTCAACTTCTCTTGGCCCGAAGAACCAAGGGCCCCGGGGGCTGGGTGATCTTTCACGTGTG	540
Y	559	gctcctcgtcgtacacctgtgtaagaggltcalcatcaaalacagagtggtcgtatcatcaacctg	618
Db	541	GCTGTGCTGTGCTGACCCCTGGTAGAGGTGCATTCATTCATTCATTCATTCATTCATTCATTC	600
Y	619	gttcggggcgaagtgtgcgaagggcgccctcaagggaacacgaacgcgcaggtc-gggccgtggc	677
Db	601	GTTCGGGGCAGTGGGAGAGGGCGCCCTCAGGGCAACAGCAGCGAGCGTGGGGCGGTGGC	660
Y	678	ctccccctgtgcacacgcgaacatggaacttgtcaatggcaactcatcaatgatactgtct	737
Db	661	CTCCCCCTGTGCGCATTCGCCCATCATGTGACTTTGTCTATGTGCGACATCTAGTATGCTGT	720
Y	738	gctcgtgggtgcctctcctcgtggggcgctggccgcgcctgtgtgtgcgctacaagcgtgtgcg	797
Db	721	GCTCGTGGGTGCTTCCTCGGGGGCGCTGGCCGCCCTGTGTGGCGCGCTCAACACGCTGGCG	780
Y	798	taagataggggtgtctgtgtgtcctctacacacgcaacctccgtgtgcacatatgtgtgtg	857
Db	791	TAACATGGGGGCTTTGTGTCTCTCACCAACACACCTCCGTTGTCATATGGGTGTGTG	840
Y	858	gacgcacatgtactactacgcgaacaa	884
Db	841	GATCGTCATGTATCTTACTTACGCGCAACAA	867
RESULT	6		
LOCUS	B1916185	905 bp	linear EST 17-OCT-2001
DEFINITION	603178072P1 NIH_MGC_121	Homo sapiens cDNA	clone IMAGE:5242600 5',
ACCESSION	B1916185		
VERSION	B1916185.1	GI:16199419	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 905)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strusberg, Ph.D.		
	Email: cga@remail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	Plate: L16M1611	row: O	column: 17
	High quality sequence stop: 883.		
FEATURES	Location/Qualifiers		
source	1..905		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5242600"		
	/gene="l1b="NIH_MGC_121"		
	/lab_host="DH10B"		
	/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;		
	Site_2: EcoRV (destroyed); RNA source anonymous pool of 3		
	fetal brains, female age 20 weeks, female age 24 weeks,		
	and male age 26 weeks. Library is oligo-dT primed and		
	directionally cloned (EcoRV site is destroyed upon		
	cloning). Average insert size 1.7 kb, insert size range		
	0.7-3.5 kb. Library is normalized and enriched for		
	full-length clones and was constructed by C. Gruber		
	(Invitrogen). Research Genetics tracking code 017. Note:		

BASE COUNT	135 a	292 c	272 g	205 t	1 others
ORIGIN	This is a NIH_MGC Library."				
Query Match	46.5%; Score 845.2; DB 10; Length 905;				
Best Local Similarity	97.8%; Pred. No. 8e-169;				
Matches 878:	Conservative	0;	Mismatches	18;	Indels 2; Gaps 2.
Qy	10	ccctaccaccgcygaagatgacgagtcgctcagcctcgtgagggagcccaacagagcctgac	69		
Db	1	ccctcaccgagccgggaagatgacgagtcgctcagcctcgtgagggagcccaacagagcctgac	60		
Qy	70	ctggagagccagagatgagccatcccaaaagccttgatgagtgctctgggagctctctctc	129		
Db	61	ctggagagccagagatgacccatcccaaaagccttgatgagtgctctgggagctctctctc	120		
Qy	130	ctgttcceccagggagcctgagcccaaggccatgctcccaaccgagctgacgcaagagcctcaac	189		
Db	121	ctgttcceccagggagcctgagcccaaggccatgctcccaaccgagctgacgcaagagcctcaac	180		
Qy	190	cccccgtacctacaacctgctgacgcgtctcgtgggagcgtgagcagctcctgagggcgtg	249		
Db	181	ccccgtacctacaacctgctgacgcgtctcgtgggagcgtgagcagctcctgagggcgtg	240		
Qy	250	gctggggcgagagatgtgccaagcttctgtcacaacatcaccctggttgagccagctccccc	309		
Db	241	gctggggcgagagatgtgccaagcttctgtcacaacatcaccctggttgagccagctccccc	300		
Qy	310	ttgtgcaagagacacaaagaaacgagagcctgcttgggagaccagatattctctctctggg	369		
Db	301	ttgtgcaagagacacacaaagaaacgagagcctgcttgggagaccagatattctctctctggg	360		
Qy	370	accctggagcctctctcgtcctgctgtgtgctgtgtgagaaacccgagctctcctcagctgt	429		
Db	361	accctggagcctctctcgtcctgctgtgtgctgtgtgagaaacccgagctctcctcagctgt	420		
Qy	430	gacctcagagcctctcctcttgagggtctgtgtcgcaatctctctctctctctgagagct	489		
Db	421	gacctcagagcctctcctcttgagggtctgtgtcgcaatctctctctctctctgagagct	480		
Qy	490	caagcctcttgccctaaactctcgtcccgaaagaaacaaagggcccgagagctgggtgac	549		
Db	481	caagcctcttgccctaaactctcgtcccgaaagaaacaaagggcccgagagctgggtgac	540		
Qy	550	ttcaactgtgagctctgctgctgacacctgtgtagagtgcatcatcaatacagagtgagctgac	609		
Db	541	ttcaactgtgagctctgctgctgacacctgtgtagagtgcatcatcaatacagagtgagctgac	600		
Qy	610	atcaacctggtctcgaggagctgagcgaaggcgagcctcagggcaacagcagcgaagcttgg	669		
Db	601	atcaacctggtctcgaggagctgagcgaaggcgagcctcagggcaacagcagcgaagcttgg	660		
Qy	670	gcccgtgagcctcccccctgtgccaatgccaacatgagagctgtgcatggagctcatcagctc	729		
Db	661	gcccgtgagcctcccccctgtgccaatgccaacatgagagctgtgcatggagctcatcagctc	720		
Qy	730	atgctgagctgagctgagctcctcctcgtggggagcctgagccgagcctgtgtgagcagctacaag	789		
Db	721	atgctgagctgagctgagctcctcctcgtggggagcctgagccgagcctgtgtgagcagctacaag	780		
Qy	790	cgctggcgtaagcaltgaggtctctgtgtcctcaccacaagccactcggttgacataagg	849		
Db	781	cgctggcgtaagcaltgaggtctctgtgtcctcaccacaagccactcggttgacataagg	839		
Qy	850	gtggtggtgagctgcatgtatatactcaagcaacagcagcaacaaagctccaccctggg	907		
Db	840	gtggtggtgagctgcatgtatatactcaagcaacagcagcaacaaagctccaccctggg	896		

```

DEFINITION AL546406.LTI.NFL006.PL2 Homo sapiens cDNA clone CS0D1031YJ06.5
prime, mRNA sequence.
ACCESSION AL546406
VERSION AL546406.1 GI:12879488
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 851)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 851
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CS0D1031YJ06"
/clone="LTI.NFL006.PL2"
/tissue_type="placenta"
/note="Vector: PCWVSPORT 6; Site: 1; NotI: 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCWVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 127 a 281 c 245 g 196 t 2 others
ORIGIN

Query Match 46.0%; Score 837.6; DB 9; Length 851;
Best Local Similarity 99.5%; Pred. No. 3.2e-157;
Matches 848; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 176 gccaaagcccaaacccctgaactacaaacccgtgtgacccgtcggggcgctggggcattcg 235
|
Db 1 gccaaagccctaacccctgaactacaaacccgtgtgacccgtcggggcgctggggcattcg 60

QY 236 tccctgagagccgtgtgctgagggcgagcatgttcacacagttgtgtcaccatcactctg 295
|
Db 61 TCCCTGAGAGCCGTGTGCTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120

QY 296 tggcagacccctcccttgtgtcagagcaacaaagagagcctgtcgggagccaggtat 355
|
Db 121 TGCGCACGCTCCCTTGTGTGAGAGACCAAGAAACGAGCGCTGTGGGAGCCAGGAT 180

QY 356 tcttcctctggggagacccctgtgctcttctgacctgtgtgtgtgtgtgtgtgtgtgt 415
|
Db 181 TCTTCTCTCTCTGAGGACCTTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240

QY 416 actctccacactgtgacctcgtcgtcgtcctctcttgggggtgtctgtcgcacatcgtct 475
|
Db 241 ACTTCTCTACACTGTGCTCTGTGGGCGGCTTCCCTTGGGGTCTGTGTGCCCATCTGCTTCT 300

QY 476 ctgtctgtgagctcagctcttgcacctcaactcctcgtgcccggagaagaacacggagccc 535
|
Db 301 CTTGTCTGGGCGGCTACAGCTTTTGCCTCAACTTCTGTGGGCCCGGAGAACACGAGGCCCC 360

QY 536 gggagctggatgtatcttactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 595
|
Db 361 GGGGCTGGGTATCTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420

QY 596 caagatgtgcatcatcacctgtgtcggggcagtgccgagggcgagccctcagaagcaaca 655
|
Db 421 CAGAGTGGCTATCATCACCTGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480

```

```

QY 656 gcaagcgagctggcgctgtgacctccctctgtgcatcgcgaacatgactgtgtcatgg 715
|
Db 481 GCAGCGCAGGCTGGGGCGCTGGCTCCCTCCCTGTGCGCATCCGCAACATGACTTTGTCTAG 540

QY 716 cactcatctacgtatctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 775
|
Db 541 CACTCATCTACGTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600

QY 776 gtggcgctacaagcgctgagcgtgaagcatggggtgtgtgtgtgtgtgtgtgtgtgtgt 835
|
Db 601 GTGGCGGCTCAACAGCGCTGGGCGTAA-SATGGGGCTTTGTGCTCTTCACACACAGCCACT 659

QY 836 ccgttgcataatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 895
|
Db 660 CCGTTGCCATATGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 719

QY 896 gtccacactggatgaccccaacgctgtgacatgcctcgcgcctccgccaatgctgtgctcg 955
|
Db 720 GTCCACACTGGGATGACCCGCGCTGGGCGATCGGCGCTGTGCCAATGCCCTGGGCGTTTCG 779

QY 956 tccctctctacgtatcccgaggtctcccaagtgaccgaagtcagccagcagaagcct 1015
|
Db 780 TCCCTCTACGTATCTCCCGAGGCTTCGCCAGTGACCAAGTYACGCCAGCAAGCAAGCT 839

QY 1016 accagggggagca 1027
|
Db 840 ACCAGGGGGAGCA 851

RESULT 8 842 bp mRNA linear EST 16-FEB-2001
LOCUS AL571754/c
DEFINITION AL571754.LTI.NFL006.PL2 Homo sapiens cDNA clone CS0D1031YJ06.3
prime, mRNA sequence.
ACCESSION AL571754
VERSION AL571754.1 GI:12929365
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 842)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1. 842
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1031YJ06"
/clone_lib="LTI.NFL006.PL2"
/tissue_type="placenta"
/note="Vector: PCWVSPORT 6; Site: 1; NotI: 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the PCWVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 147 a 248 c 276 g 166 t 5 others
ORIGIN

Query Match 43.6%; Score 792.6; DB 9; Length 842;
Best Local Similarity 98.1%; Pred. No. 1.1e-157;

```

Matches	828: Conservative	4: Mismatches	9: Indels	3: Gaps	3:
QY	871	acttaacggaacagcaacagcaacagctccaccatggaatgacccaagctgacatcgcc	930		
Db	842	ACTTACCGGCAACAAACAGACGACMACAGTCCACCTGGATGATACCCACGCTGGCCATCGCC	783		
QY	931	ctcgccgcgaatagctctgagccttgctccctctctgaatcccccgaagctccccaagt	990		
Db	782	CTCGCTGCCAAATGCCCT-GGCCCTTGCTCTCTCTACGTATATCCCGA-GTCTCCAGGTG	725		
QY	991	accaagctccagcccgagagaaagactccagagggagacatgtaacccaccggggctgggc	1050		
Db	724	ACCAAGTCCAGGCCACAGAGCAAAAGCTACCAAGGGGSATGTATACCCACCCGGGGCTGGGC	665		
QY	1051	tatgagaccatctcgtaaagagcagaagaggtcagagcatgtctgtgagagaacaagccttt	1110		
Db	664	TATGAGACCATCTCTGAAGAGCAGAAAGGCTAGACATATGTTGCTGAGAACAAAGCCCTTT	605		
QY	1111	tccatagatgagccggtctgacgactaagaggccggtgctacatcaacgagtgtaaatgg	1170		
Db	604	TCCATAGATGAGACCCGCTGCACCTAAAGGCCGGGTGTACCCCTAACAGGGGTACATGGG	545		
QY	1171	cagctgcgcgaacagctgtgtacagcccaacttgatctggccctgatagcacaagtctcgctcc	1230		
Db	544	CAGCTGCTBACCAAGTGTGTACCAAGCCCACTGATGATGCCCCCTGATGCACAAAGTTCCGTTCC	485		
QY	1231	gaagagagcttacaatcatcctcccaagggcaccacgccaacagccaggtgaatggcagt	1290		
Db	484	GAAAGCGCTTACGACCTCATCTCTCCACAGGGCCACGCCACACAGCAGGATGATGGGCAGT	425		
QY	1291	gccaaactcgaccctgcgggctgtaaaacatgtatctcgcccaagaccacaagcgggccaca	1350		
Db	424	GCCCACTCGCA-CTGCGGGCTGAAAGACATGTACTGGGCCACAGCCMCAGCGGCGCACAA	366		
QY	1351	ccggcgaaagagcggcaagaactctgaagttctttagaaaccttaagttgtgtgagactgagtc	1410		
Db	366	CCGCCGAAAGACGCAAGACACTCTCAAGTCTTTTAAACACCTTACGTGTGGAGACTGAGTIC	306		
QY	1411	agcggatggcgaagagagcggctgcggaatttgaggaaagcccttgaagacacttgscggcgca	1470		
Db	306	ACCGGTGGCGAAGAGAGGGGGGTGGGATTTGGGAGAGGCCCTTGAGGACCTGGCCCCGGGGA	246		
QY	1471	agggaactctcagctctcctctccctctgcagcggcccaagaacaatgtgccagaatgtgg	1530		
Db	246	AGGGACTCTCCAGGCTCCTCCCTCCGCCCTGGGCAAGGCCAACATGTGCCAGATGTGG	186		
QY	1531	aagggcctcccccctctctgacagtgatttgagggtgtcaagggtgtgtcccaaccacatctctc	1590		
Db	186	AAGGCTCCTCCCTCTCTGCGCAGTGTGTGGTGGGTGTCAATGGGTGTGCCACCCACTCTCTC	126		
QY	1591	agtgcttctgtagctcgagagagccaaccccaagcctctctgcagaatcactctggcggtcac	1650		
Db	126	ACTGTTGTGGAGTGTGAGAGAGCCAAACCCAGCTCTCTGCACAGATCACCTCGCGGTAC	66		
QY	1651	actccagcccaaatagtgctctctcgagggtgtgctgagcgacgcgtatgttctcttgaga	1710		
Db	66	ACTCCAGCCAAATAGTGTCTCTGGGGGTGTGTGCTGGCGCCTATATGTTTCTCTGGAGA	6		
QY	1711	tccc 1714			
Db	5	TTCC 2			
RESULT	9				
LOCUS	BG829477	956 bp	mRNA	linear	EST 22-MAY-2001
DEFINITION	60276377441 NIH_MGC_42 Homo sapiens CDNA clone IMAGE:489018 5',				
ACCESSION	BG829477				
VERSION	BG829477.1	GI:14177064			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 956)	NIH-MGC http://mgc.nci.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	Unpublished (1999)
			Contact: Robert Strausberg, Ph.D.	Email: cgabds-remail.nih.gov
			Tissue Procurement: ATCC	
			cDNA Library Preparation: Ling Hong/Rubin Laboratory	
			cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
			DNA Sequencing by: Incyte Genomics, Inc.	
			Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:	
			http://image.llnl.gov	
			Plate: LILCM1790	row: c column: 19
			High quality sequence stop: 770.	
			Location/Qualifiers:	
			1..956	
			/organism="Homo sapiens"	
			/db_xref="taxon:9606"	
			/clone="IMAGE:4899018"	
			/clone.lib="NIH-MGC.42"	
			/tissue.type="epithelioid carcinoma cell line"	
			/lab.host="DH10B (phage-resistant)"	
			/note="Organ: pancreas; Vector: pOTB7; Site.1: XhoI; Site.2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library. 1"	
BASE COUNT	147 a	300 c	293 g	216 t
ORIGIN				
Query Match	42.6%;	Score 775.2;	DB 10;	Length 956;
Best Local Similarity	96.1%;	Pred. No. 5.5e-154;		
Matches 848;	Conservative	0;	Mismatches 28;	Indels 6;
				Gaps 5
QY	17	caagcggaaagtaagtcagtcgctcagccttgaggaggaaccaacagagctgagcctggag	76	
DB	2	CAGCGGGAAAGTACGATCGATCGCTCAACCTCGAGAGGACCCACAGACGCTGCGAGAG	61	
QY	77	caagatgagccttcaacaagccttgatgatgttgccttgagactgctctctccagtcttc	136	
DB	62	CCAGGATGGCCATCCCAAGAGCCTTGATGATGCTTGAGAGCTGCTCTCTCTCTCTCTTC	121	
QY	137	caagggcctcgagcccaaggacatgctccaccacccgagctgcagcaagcctcaaacccctgt	196	
DB	122	CAGGGGCGTGGGGCCAGGGCCAGTGTCCACCCGGCGTGCAGCCAGGCGTCAACCCCTGT	181	
QY	197	actacaacctgtgtgacgcgctctggggcgctggggcaatcgctctgagagccgtgctggag	256	
DB	182	ACTAACACTGTGTGATGCCCTGTGGGGGCGATCTCTCTGGAGGCGCTGTGCTGGGG	241	
QY	257	cgggatgttcaaccagcttgtgtccacatatcctgttgccagagccctccctctgtgc	316	
DB	242	CGGGCATTTGTCAACACGTTTGTGCTACCATCATCTCTGGTGGCCACACCTCCCTTTGTGC	301	
QY	317	aggacaaccaagaacgagacgctgcctgaggaaccagatattcttctctcttgaggaccctg	376	
DB	302	AGGACACCAAGAACCGAGACCTGTGGGAGACCCAGGTATTCTTCTTCTTGGGGACCTGG	361	
QY	377	gacctcttcgctctgttctgtcctgtgtgaggaagccggaacttccacactgtgctctc	436	
DB	362	GGCTCTTGCTCGTGGTTGGCTGTGTGTGAAGGCCACATTCACACTGTGCTCTTC	421	
QY	437	ggagctctctcttgggggtctgttcgcacatctgtctctctgtcttgaggagctaacgtc	496	
DB	422	GGCGCTTCCTCTTTGGGGTGTCTTGCCACATCTGCTTCTCTTGTCTGGGGGCGTCAAGTCT	481	

Oy	497	ttccctcaacttcctggtcccggaagaacccagggccccgggctcgggtgtaattccactg	556
Db	482	ttgcccttaactccttgaccgccgaagaacacacggccccgggctgggtgtaattccactg	541
Oy	557	tgcctctgcctgaaccctgtatagatgcatcatcatatagaagatgctgtaattccacc	616
Db	542	ttggcctctctgtatgaccttgatgaggtcatcatatatacagaagtggtgatcatccacc	601
Oy	617	tggltcggggaatgtgcgaaggcggtccctcagggcaacagcagcgcaagcttggccgttg	676
Db	602	tggttcggggcagtggtggcagggcgcg-cctcagaggcaacagcagcgcaagcttggccgttg	660
Oy	677	cctcccccctgtgcatatgcacatagaactttgtatagactcatctaacgtcatgc-tg	735
Db	661	ccctcccccctgtggccatgcaccacaacatggacattgtgatggcactcatcattacgtcatgcttg	720
Oy	736	ctgcctgcctgagttgccttcctctgaggagccgcgcgcctgtgtggccgtctaacagcgctgg	795
Db	721	ctgcctgcctgggggtgctttcctctggggggccttgcccccctttgtgtggcccgtaacaaagccttgcc	780
Oy	796	cgtaagcatggggcttcttgt-gctctctacccacagccacactccgttgtcatatgggt-gg	853
Db	781	gtaagacatggggcttctgtgggtgctcctcaacaagagacacactccgttgtcatatgggtggg	840
Oy	854	tgtgtaactgcgtcatg--tatacttaaggcaacaagcagcaca	893
Db	841	tgtggatcgtcatggtatttactttacgggaacaaacacacaca	882

BASE COUNT	148 a	250 c	269 g	175 t	4 others
ORIGIN					

Query Match	41.8%;	Score 761.2;	DB 9;	Length 846;
Best Local Similarity	97.8%;	Pred. No. 4.9e-151;		

[illegible]

RESULT	11				
BI769602					
LOCUS	BI769602	873 bp	mRNA	linear	EST 25-SEP-2001
DEFINITION	603054483c1 NIH_MGC_122	Homo sapiens	CDNA	clone IMAGE:5204373	5',
	mRNA sequence.				
ACCESSION	BI769602				
VERSION	BI769602.1	GI:15761180			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 873)
NIH-MGC http://mgc.ncl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LRAM1512 row: f column: 22
High quality sequence start: 3
High quality sequence stop: 821.
Location/Qualifiers

FEATURES
source

1. .873
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5204373"
/clone_lib="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6; site_1: NotI; site_2: EcoRV (destroyed); RNA source anonymous pool of 24 week female lung, 16 week female spleen, and 20-22 week male spleens. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 026. Note: this is a NIH-MGC Library."
BASE COUNT 126 a 285 c 261 g 201 t
ORIGIN

Query Match 41.1%; Score 747.8; DB 10; Length 873;
Best Local Similarity 95.2%; Pred. No. 3,4e-148;
Matches 837; Conservative 0; Mismatches 32; Indels 10; Gaps 6;
Qy 40 cagcctgtagggagcccaagcagcctgctgtaggagccagatgcatcacacaagcc 99
Db 1 CAGCCTGAGGAGGACCCACGAGAGCTGG-CTGGAGCCAGAGATGCCATCCACAACCC 59
Qy 100 ttgtgtatgtgctgtaggagcgtctctcttctgttccaggggctggtcccaaggcat 159
Db 60 TTGTGTATGTGCTGTGGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 117
Qy 160 gtcccaaccgctgtaggagcagcagcctcaaccctgtactaaccctgtgtagcgcct 219
Db 118 GTCCACACCGGCTGAGCCAGGCTCAACCCCTGTACTAACCTGTGTAGCCGCTCT 177
Qy 220 ggggcgtgtaggagcagcagcgtgtaggagcgtgtaggagcagcagcagcagc 279
Db 178 GGGGCGTGGGCGATGCTCTGAGAGCGCTGGCTGGGGCGGCGATGTACACAGTTTGTG 237
Qy 280 ctcaacatcatctgtaggagcagcctcccttctgtcagagaacccaagaagagagc 339
Db 238 CTCACATCATCTGTGTGGAGCTCCCTCTTGTGTGAGAGACCAAGAAAGAGAGCTG 297
Qy 340 ctggagagaccagatattctctctctgtaggagccctgtagcctctctgctgttggc 399
Db 298 CTGGGAGACCCAGGATATCTCTCTGTGGGAGCCCGGCGCTCTCTCTCTCTCTCTCT 357
Qy 400 tgtgtgtaggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 459
Db 358 TGTGTGTGAGAGCCCGAGATTTCTCACCTGTGCTGTGGCGCTTCTTGTGGGTTTCTG 417
Qy 460 ttgcacatctctctcttctgtcgtgagcagcagcagcagcagcagcagcagcagc 519
Db 460 TTTTCT

Db 418 TTGGCATCTGCTCTCTCTCTGTGCTGGCGGCTCAGCTCTTTGCCCTCACTTCGCCCCG 477
Qy 520 aagacacagggccccgggggctgggtgtagtcttcaactgtgagcctgtcagcctgtgta 579
Db 478 AAGAACACAGGCCCCGGGGGCTGGGTGATCTTCACTGTGGCTGTCTGACCTGGTA 537
Qy 580 gaggatcatcatatagcagctgagctgtagtcaaccctggtctgggagcagctgagcagc 639
Db 538 GAGGTATCATATATAGAGAGTGGCTGATCATACACCTGTGGGCGGAGTGGGAGGCG 597
Qy 640 ggcctcagggcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 699
Db 598 GG-CCTCAGGGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 656
Qy 700 atgagc-cttgcattgagcattcatcagtcagtcagtcagtcagtcagtcagtcagtcag 758
Db 657 ATGACCTTTTGTGATGACATCATATACCTATATGCTGCTGCTGCTGCTGCTGCTGCTG 716
Qy 759 ggcctgagccgcccctgtgtgagcagcagcagcagcagcagcagcagcagcagcagcag 818
Db 717 GGCCTGGGCGGCGGCTGTGTGGCGGCTACCAAGCGCTGGGCTAAGCATGGGGCTTTGTGCT 776
Qy 819 cctcaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 876
Db 777 CCTCACACAGCCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 836
Qy 877 ggcacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 915
Db 837 GGCAAAAGCAGCCACACA---GTCCACCTGGGATGACCCC 872

RESULT 12

BI834606 820 bp mRNA Linear EST 04-OCT-2001
BI834606
LOCUS 603089872P1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5228800 5',
DEFINITION mRNA sequence.
ACCESSION BI834606
VERSION BI834606.1 GI:15946156
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 820)
NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LRAM1512 row: p column: 17
High quality sequence stop: 804.
Location/Qualifiers

FEATURES

source

1. .820
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5228800"
/clone_lib="NIH_MGC_120"
/lab_host="DH10B"
/note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; site_1: NotI; site_2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics

BASE COUNT 118 a 264 c 250 g 188 t
ORIGIN

Query Match 40.9%; Score 744.4; DB 10; Length 820;
Best Local Similarity 97.3%; Pred. No. 1.7e-147;
Matches 800; Conservative 0; Mismatches 16; Indels 6; Gaps 4;

QY 38 ctacagcctgagagagaccacagagccttgccctggagagccagatggtccatccacaag 97
DB 1 CTACAGCCTGAGAGGGA-CCAAACCAAGAGCTGGCCCTGGAGCCAGATGCGCATCCACAAG 59
QY 98 ccttgagatgctgctggagactgctctctctctctctctctctctctctctctctctct 157
DB 60 CTTGGTGTGATGTGCTTGGAGACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 119
QY 158 atgtccaccacgctgtgacgacgaagcctcaaccctgtactacaacctgtgtacgct 217
DB 120 ATGTCCACCCGCGCTGAGAGCAAGGCTCAACCCCTGTACTACACCTGTGTACCGCT 179
QY 218 ctggagcgctggagagatcgtctgagagcgctgtgctggagcgagcattgtacacgcttg 277
DB 180 CTGGGGCGTGGGCGATGCTCTGGAGGCGCTGGGGCGGCGCATGTCTACACAGTTTG 239
QY 278 tgcacacatcatcctgtgtgagcagcctccctctgtgcaagagacacaaagaaagagcc 337
DB 240 TGTCTACATCATCTCTGTGTGGCCAGCCTCCCTTTGTGACAGACCAAGAAAGGAGCC 299
QY 338 tgcctggagaccagatct 397
DB 300 TGCTGGGGACCCAGGATATCT 359
QY 398 ccgtgtgtgtgagagccgagactctcaacgtgtgctctctctctctctctctctctct 457
DB 360 CCTGTGTGTGTGAACCCGAGCTTCTCCACCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCT 419
QY 458 tgttcgcacatcgtct 517
DB 420 TGTTCGCTCATCT 479
QY 518 ggaagaacacagcgcccgagcgctgggtgactctcaactgtgtgctctctctctctct 577
DB 480 GGAAGAACAAGCGGCCCGGGGCTGGGATCTTCACTGTGTGCTCTGTCTGTACCCCTGG 539
QY 578 tagaggtcatcatcaatacagagatgctgtatcatcaacctgtgtctggggcagtgag 637
DB 540 TAGAGGTTCATCATCAATACAGAGTGGCTGATCATCACCTGTGCTGGGGCAGTGAGG 599
QY 638 ggggcctcaaggggaaagagcagcgctgtggcgctgtgctctccctctgttccatgcca 697
DB 600 GCGGACCTCAAGGCAAGGAGCGAGCGAGCGGTGGCCCTCCCTCTGTCCATCGCCA 659
QY 698 acatgagactgtcatgtgacactcatctacgtcatgtctgtctgtgtgtgtgtgtgt 757
DB 660 ACATGGAGCTTTGTGATGGCACTCATACGTCACTGTCTGTCTGTGGGTCTCTTCCTGG 719
QY 758 gggcctgtggcccgctgt 817
DB 720 GGGCCTGTGGCCCG-CCGTGTGTGGAGC-TACAAAGCGCTGGCGTAAGCTGTAAAGTGC 777
QY 818 tctctaccacagacagcctcgttgcacatagtgtgtgtgtgtgtgtgtgtgtgtgtgt 859
DB 778 TCCT--CACACAGGAGCTCCGTGCATATGTGATGTGTGA 816

RESULT 13
BI772730 933 bp mRNA linear EST 25-SEP-2001
LOCUS 603035260F1 NIH_MGC_122 Homo sapiens cDNA IMAGE:520285 5',
DEFINITION mRNA sequence.
ACCESSION BI772730
VERSION BI772730.1 GI:15764308

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LHAM11508 row: 9 column: 16
High quality sequence stop: 782.

FEATURES

source

Location/Qualifiers
1..933
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5202855"
/clone_1lb="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."

BASE COUNT 141 a 303 c 276 g 213 t
ORIGIN

Query Match 40.2%; Score 732; DB 10; Length 933;
Best Local Similarity 92.9%; Pred. No. 7.7e-145;
Matches 866; Conservative 0; Mismatches 55; Indels 11; Gaps 9;

QY 55 caacagagcctggcctgtggagcagatggtccatccaaagcctgtgtatgtgctgt 114
DB 3 CAACGAGAGCCTGG-CTGGAGACCAAGATGGCATCCAAAGCCTTGTGATGTGCTGTG 61
QY 115 gg-actgtct 173
DB 62 GGAAGCTGCT 119
QY 174 cagccaaagctcaacccctgtactacaacctgtgtacgctctgtggcgctgtggagat 233
DB 120 CAGCCAAAGCCTCAACCCCTGTACTACACCTGTGTACCGCTGTGGGGGTGGGAGAT 179
QY 234 cgtctctgagcgctgtgctgtggcgagcattgtacacacgttctgtctacatcatcct 293
DB 180 CGTCTGTGAGCGCGTGGCTGGGGCGGCGCATGTGCACACGCTTGTGTGTCACACATCATCT 239
QY 294 ggtggcagcctccctctctctgtgcaagacaagaagaagagcctgtcgtggagacccag 353
DB 240 GGTGGCCAGCTCCCTCTTGTGTGCAAGGACCAAGAAAGGAGCCGTGTGGGGACCCAGGT 299
QY 354 attctctctctgtggagaccctgggacctctctgtcctctgttctgctgtgtgtgtgtgtgt 413
DB 300 ATTCTCTCTTGTGGGAGCCCTGGGCGCTTCTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGT 359
QY 414 cgaactctcaacctgtgtcct 473
DB 360 CGACTTTCACACGTGTGCTCTGGGCGCTCTCTCTTGTGGGCTGTGTGTGTGTGTGTGTGTGT 419

OY	474	ctcttgcttgagggtactaagctctttgcccctaacttctctggcccggaagaacccaagggcc	533
Db	420	ctcttgcttgagggtactaagctctttgcccctaacttctctggcccggaagaacccaagggcc	479
OY	534	ccggagcttgagtgatcttcaacttgagctctgcctgcttgacccttgtagagtgatcatcaaa	593
Db	480	ccggagcttgagtgatcttcaacttgagctctgcctgcttgacccttgtagagtgatcatcaaa	539
OY	554	tacaaagtgagctgatacattcaacccttggttcggggcagtgagggcgccctccagggcaa	653
Db	540	tacaaagtgagctgatacattcaacccttggttcggggcagtgagggcgccctccagggcaa	598
OY	654	caggagcgcaagcttgagccgttgagctctcccccctgtagcatcagcaacaatgaaatttgcat	713
Db	599	caggagcgcaagcttgagccgttgagctctcccccctgtagcatcagcaacaatgaaatttgcat	657
OY	714	gggaactcatctacgtcatalagctgagctgctctggttgagtgactcct--ggggagctggccggccc	772
Db	658	gggaactcatctacgtcatalagctgagctgctctggttgagtgactcct--ggggagctggccggccc	717
OY	773	tgtgtggcccgctacaaagcgctgagcgtlaagcat--ggggctcttggtctccttcaaccaagcc	831
Db	718	tgtgtggcccgctacaaagcgctgagcgtlaagcat--ggggctcttggtctccttcaaccaagca	777
OY	832	accttcgltgccaat--gggttggttgatctgcatatgatacttaacggaacaagcaagc	889
Db	778	accttcgltgccaat--gggttggttgatctgcatatgatacttaacggaacaagcaagc	837
OY	890	acaacagatcccaacccggg--atgacccaagcgtgagcatctgcgctccggcccaatgcctg	948
Db	838	acaacagatcccaacccggg--atgacccaagcgtgagcatctgcgctccggcccaatgcctg	897
OY	949	gacctgcgtcctctctacgtcaatcccggaagt	980
Db	898	gacctgcgtcctctctacgtcaatcccggaagt	929

```

RESULT 14
BG750651
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BG750651 776 bp mRNA linear EST 15-MAY-2001
602708511F1 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:4844930 5',
RNA sequence.
BG750651
BG750651.1 GI:14061304
EST.
human.
Homo sapiens
Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 776)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
http://image.lnl.gov
Plate: LNCM1683 row: f column: 03
High quality sequence stop: 750.
Location/Qualifiers
1..776

```

```

"organism":"Homo sapiens"
"db_xref":"taxon:9606"
"clone_image":"4844930"
"clone_lib":"NH_MGC_43"
"tissue_type":"normal pigmented retinal epithelium"
"lab_host":"DH10B (phage-resistant)"
"note":"Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:

```

ECORI, cDNA made by oligo-dT priming, directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC library. | "

BASE COUNT	169 a	245 c	231 g	131 t
ORIGIN				

Query Match	40.0%	Score 728.4;	DB 10;	Length 776;
Best Local Similarity	99.0%;	Pred. NO. 4.2e-144;		
Matches 754; Conservative	0;	Mismatches 6;	Indels 2;	Gaps 2

QY	861	cgtaatgtatacttaacgcgaacaagcagcaacaacagctccacacctggatgtgaccccaagct	920
Db	2	ccctcatgtatattcttaccggcaaacacacagcaaacagctccacacctggatgtgaccccaagct	61
OY	921	ggcaatcgacccctgcgcgcgccaatgcctgggaccttgctccctctctacagtcacatccccaagt	980
Db	62	ggccatctgcgacctctgcgcgcgccaatgcctgggaccttgctccctctctacagtcacatccccaagt	121
OY	981	ctcccaagtgacccaagtctccaagcccaagcaaaagtataccaaggggagacatgtaccccaaccg	1040
Db	122	ctcccaagtgacccaagtctccaagcccaagcaaaagtataccaaggggagacatgtaccccaaccg	181
OY	1041	gggcgttgggctatgtgagccaatcccttgaagaagcagaagggctcagagcatgttcgttggagaa	1100
Db	182	gggcgttgggctatgtgagccaatcccttgaagaagcagaagggctcagagcatgttcgttggagaa	241
OY	1101	caagagcctttccatagatgaagcggcttgacgttaagagcgggtgtacacatacagcg	1160
Db	242	caagagcctttcttcattatggaatgagccgcttgacgttaagagcgggtgtacacatacagcg	301
OY	1161	gtacaaatgggcagcgtctgtgacagatgtgttlaacagcccacttgatgatgtgacctgtatgcaaa	1220
Db	302	gtacaaatgggcagcgtctgtgacagatgtgttlaacagcccacttgatgatgtgacctgtatgcaaa	361
OY	1221	agttctgtccgaagagagcttagacataatctctccaagggcgcaacggccaagccaggt	1280
Db	362	agttctgtccgaagagagcttagacataatctctccaagggcgcaacggccaagccaggt	421
OY	1281	gattgggcagtgacaaactcgacacctgcgagctgaagaacatgtaactcggcccgaagccacca	1340
Db	422	gattgggcagtgacaaactcgacacctgcgagctgaagaacatgtaactcggcccgaagccacca	481
OY	1341	ggcggccacaacccgcggaagacgagcaagaactctcaaggtcctttagaacccttaeqtltg	1400
Db	482	ggcggccacaacccgcggaagacgagcaagaactctcaaggtcctttagaacccttaeqtltg	541
OY	1401	ggaactgaugtcaagcgttggcgagaagaaggggtcggatatttggggaagggcccttaagaacctg	1460
Db	542	ggaactgaugtcaagcgttggcgagaagaaggggtcggatatttggggaagggcccttaagaacctg	600
OY	1461	ggcccgggcaagagactctccaggtctcctctccccccttgcagcgcccgacaacatgtgctc	1520
Db	601	ggcccgggcaagagactctccaggtctcctctccccccttgcagcgcccgacaacatgtgctc	660
OY	1521	ccagaatgtgaaagggcctcctcctctctgcacagtgttgggtgggtgtcaltatgggtgtcccca	1580
Db	661	ccagaatgtgaaagggcctcctcctctctgcacagtgttgggtgggtgtcaltatgggtgtcccca	719
OY	1581	cccaactcctcagtgcttgggtgggtgtcaltatgggtgtcccca	1622
Db	720	cccaactcctcagtgcttgggtgggtgtcaltatgggtgtcccca	761
RESULT 15			
LOCUS	B1822698	1052 bp	mrna linear EST_04-OCT-2001
DEFINITION	603036033p1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176977 5',		
mrna sequence.			

ACCESSION BI822698
 VERSION BI822698.1 GI:15934248
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1052)
 TITLE NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: c9apbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1AM11441 row: a column: 10
 High quality sequence start: 2
 High quality sequence stop: 736.
 Location/Qualifiers
 1..1052
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5176977"
 /clone_1b="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: This is a NIH_MGC Library."
 BASE COUNT 204 a 338 c 306 g 204 t
 ORIGIN
 Query Match 39.0%; Score 710.2; DB 10; Length 1052;
 Best Local Similarity 93.8%; Pred. No. 3.4e-140;
 Matches 806; Conservative 0; Mismatches 43; Indels 10; Gaps 6;
 Oy 45 tggaggagacccaacacgctggccttgaggacagagatggccatccacaagaagccttgt 104
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 tggagggagacccaacacgacgctggccttgaggacagagatggccatccacaagaagccttgt 60
 Oy 105 gatgtgctggagactgct 164
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 gatgtgctggagactgct 119
 Oy 165 accgggctgacgaaggctccaaacccctgtactaacaacctgtgtgacgcgtctgggac 224
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 120 accgggctgacgaaggctccaaacccctgtactaacaacctgtgtgacgcgtctgggac 179
 Oy 225 gtggggacatgctcctggagagccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 284
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 180 gtggggacatgctcctggagagccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 239
 Oy 285 catcatcctggatggcacaagctcccccctgtgtcagagacacaaagaacgagcctgtggg 344
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 240 catcatcctggatggcacaagctcccccctgtgtcagagacacaaagaacgagcctgtggg 299
 Oy 345 gacccaaggtatctgtgt 404
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 300 gaccacaggtatctgtgt 359
 Oy 405 ggtgaagcgcgactctcactgtgctctctggcgcttccctcttggggtctgtctgc 464

||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 360 ggtgaagcgcgactctcactgtgctcctggggcttccctcatatgggggttctgttgc 419
 Oy 465 catctgtctctctgtctgtgcggtcagctctcttgcctcaactctctgtgccggaaga 524
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 420 catctgtctctctgtctgtgcggtcagctctcttgcctcaactctctgtgccggaaga 479
 Oy 525 ccacgggcccggggcttgggtgactctgactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 584
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 480 ccacgggcccggggcttgggtgactctgactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 539
 Oy 585 catcatcaatacagaatgtgctgactcatcaccctgtgtcgggacagtgagggcgagcc 644
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 540 catcatcaatacagaatgtgctgactcatcaccctgtgtcgggacagtgagggcgagcc 599
 Oy 645 tcaaggacaa-cagcaagcaggtctgggcgtgtgctcccccctgtgtcgaatgcaaa-catg 702
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 600 tcaggacaaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 659
 Oy 703 gactttgcatggagactcatcatcagctcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 762
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 660 gactttgcatggagactcatcatcagctcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 719
 Oy 763 tggccggcctgtgtgtgcgctacaagcgctgcg-----taagcatggggtcttctgtgc 817
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 720 tggccggcctgtgtgtgcgctacaagcgctgcg-----taagcatggggtcttctgtgc 779
 Oy 818 tctcaaccac-agccaactccgttgcacat-aggatgtgtgtgtgtgtgtgtgtgtgtgtgt 875
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 780 tctcaaccac-agccaactccgttgcacat-aggatgtgtgtgtgtgtgtgtgtgtgtgtgt 839
 Oy 876 cggcaacagcagcagcaac 894
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 840 acaggaacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 858

Search completed: September 22, 2002, 16:13:13
 Job time: 5936 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 14:39:52 : Search time 75.36 Seconds
(without alignments)
5928.972 Million cell updates/sec

Title: US-09-895-686-7

Perfect score: 1819
Sequence: 1 cgcctcgagccctaccagc.....ctttatcttaacttaaaaa 1819

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing First 45 summaries

Database : Issued Patents.NA:*
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCRTOS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	191.6	10.5	2484	4	US-09-276-531-46 Sequence 46, Appl
2	187.4	10.3	1212	3	US-09-188-930-249 Sequence 249, App
3	118.8	6.5	311	3	US-09-188-930-4 Sequence 4, Appl
4	52.4	2.9	7218	1	US-08-232-463-14 Sequence 14, Appl
5	46.2	2.5	1035	3	US-08-733-837B-1 Sequence 1, Appl
6	45.4	2.5	1931	2	US-09-130-114-2 Sequence 2, Appl
7	45.2	2.5	6048	4	US-09-634-920-3 Sequence 3, Appl
8	43.6	2.4	4411529	4	US-09-103-840A-1 Sequence 1, Appl
9	41.6	2.3	1597	2	US-08-724-974A-1 Sequence 1, Appl
10	40.8	2.2	3384	2	US-08-687-289A-1 Sequence 1, Appl
11	40.8	2.2	3809	1	US-08-485-588-3 Sequence 3, Appl
12	40.8	2.2	3809	1	US-08-484-565-3 Sequence 3, Appl
13	40.8	2.2	3809	2	US-08-480-751-3 Sequence 3, Appl
14	40.8	2.2	3809	2	US-08-943-986-3 Sequence 3, Appl
15	40.8	2.2	3809	3	US-08-353-784-3 Sequence 3, Appl
16	40.8	2.2	3809	3	US-08-484-719B-3 Sequence 3, Appl
17	40.8	2.2	3809	4	US-08-546-998-2 Sequence 2, Appl
18	40.8	2.2	3809	4	US-08-484-159-3 Sequence 3, Appl
19	40.8	2.2	4465	1	US-08-180-195-1 Sequence 1, Appl
20	40.8	2.2	4465	1	US-08-477-329-1 Sequence 1, Appl
21	40.8	2.2	4465	2	US-08-475-458-1 Sequence 1, Appl
22	40.8	2.2	4465	3	US-08-980-400-1 Sequence 1, Appl
23	40.8	2.2	4465	4	US-09-583-459A-1 Sequence 1, Appl
24	40.8	2.2	4465	4	US-09-583-210-1 Sequence 1, Appl
25	40.8	2.2	4465	4	US-09-583-449A-1 Sequence 1, Appl
26	40.8	2.2	5006	1	US-08-485-588-2 Sequence 2, Appl
27	40.8	2.2	5006	1	US-08-484-565-2 Sequence 2, Appl

28	40.8	2.2	5006	2	US-08-480-751-2 Sequence 2, Appl
29	40.8	2.2	5006	2	US-08-943-986-2 Sequence 2, Appl
30	40.8	2.2	5006	3	US-08-353-78A-2 Sequence 2, Appl
31	40.8	2.2	5006	3	US-08-484-719B-2 Sequence 2, Appl
32	40.8	2.2	5006	4	US-08-546-998-1 Sequence 1, Appl
33	40.8	2.2	5006	4	US-08-484-159-2 Sequence 2, Appl
34	40.8	2.2	5427	1	US-08-168-917-1 Sequence 1, Appl
35	40.8	2.2	5427	2	US-08-460-510-1 Sequence 1, Appl
36	40.8	2.2	5427	2	US-08-460-490-1 Sequence 1, Appl
37	40.8	2.2	5427	3	US-08-462-728-3 Sequence 3, Appl
38	40.8	2.2	5427	5	PCR-US92-00730-1 Sequence 1, Appl
39	40.8	2.2	5427	5	PCR-US92-00862-1 Sequence 1, Appl
40	40.6	2.2	289	4	US-09-007-005-17 Sequence 17, Appl
41	40.6	2.2	289	4	US-08-244-796-17 Sequence 17, Appl
42	40.6	2.2	3489	2	US-08-728-323A-1 Sequence 1, Appl
43	40.6	2.2	32207	2	US-08-770-379-20 Sequence 20, Appl
44	40.6	2.2	32207	4	US-08-757-669A-20 Sequence 20, Appl
45	40.6	2.2	32207	4	US-09-230-371A-20 Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-276-531-46
Sequence 46, Application US/09276531
Patent No. 6183968
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Yue, Henry
APPLICANT: Reddy, Koopa
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
NUMBER OF SEQUENCES: 134
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/276, 531
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/079, 677
FILING DATE: March 27, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Lynn E. Murry, Ph.D.
REGISTRATION NUMBER: 42,918
REFERENCE/DOCKET NUMBER: PA-0008 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 2484 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYNORAT05


```

? PRIOR FILING DATE: 1999-08-09
? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 3
? LENGTH: 6048
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(6048)
?
? US-09-634-920-3

```

Query Match	2.5%;	Score 45.2;	DB 4;	length 6048;
Best Local Similarity	46.3%;	Pred. No. 0.12;		
Matches 149;	Conservative	0;	Mismatches 173;	Indels 0;
			Gaps	0;

[illegible]

```

RESULT      8
US-09-103-840A-1
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007_00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

Query Match	2.4%	Score 43.6	DB 4	length 4411529;
Best Local Similarity	44.6%	Pred. NO. 2		
Matches	172	Conservative	0	Mismatches 214; Indels 0; Gaps 0;
OY	609	cataccctggttcggagcgatgcgaaggcgcgcctcatggacaaagcgcagctg	668	
Db	1863322	cataaccgagcgacgcgcccgggtctgcgcgcgcgcgcgcgcgcgcgcgcgcgttac	1863381	

[illegible]

RESULT 9
 US-08-724-974A-1
 Sequence 1, Application US/08724974A
 Patent No. 5912335
 GENERAL INFORMATION:
 APPLICANT: Derek J. Bergsma, Catherine E. Ellis
 TITLE OF INVENTION: A No. 5912335e1 G-Protein Coupled Receptor
 TITLE OF INVENTION: HUVCT36
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham Corporation
 STREET: 709 Swedeland Road, P.O. Box 1539
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19406-0939
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
 COMPUTER: IBM 486
 OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/724,974A
 FILING DATE: October 3, 1996
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: William T. Han
 REGISTRATION NUMBER: 34,344
 REFERENCE/DOCKET NUMBER: ATG50022
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610 270 5024
 TELEFAX: 610 270 5090
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1597
 TYPE: Nucleic Acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 ANTI-SENSE: No
 US-08-724-974A-1

Query Match	2.38;	Score 41.6;	DB 2;	Length 1597;
Best Local Similarity	44.48;	Pred. No. 0.57;		

	Matches	167;	Conservative	0;	Mismatches	209;	Indels	0;	Gaps	0;
QY	675	gqccctccccctgtgcacatcgccaataygacttctatgcaactcatctaatglatgct	734							
Db	878	ggcAATGGAGCGCGGCGCATTCACATCTACCGCTCTCTGGGGGCTTCCTTCCCATCTGG	937							
QY	735	gctgcctgtctgggtgccttcctctgggggacctgagccgacctgtatgagcgctacaagcgctg	794							
Db	938	CCTGCTGCTGGCGCTCTTACCAAGGCGATCTCGGGCGCGCGGAGCGACCGACCCA	997							
QY	795	gcgtaagcaatggygcttctgttgctctctcaaccaagccacctcgltgcataatggytgat	854							
Db	998	GAAGAGCGCGAAGGACGACGATCCACAGCGGCTGCTGCTCAAGACCGTGTCATCTTCTGGC	1057							
QY	855	gtggaatgctcatgtatctacttaactaaggaacaaagagcaacaaagatcccaacttgatgaccc	914							
Db	1058	CTGCTTCTCTGCCCTTACACACAGTGTTGCTGCTGCTGGCGCAAGCTGTGGAGGCCAGCTGCCA	1117							
QY	915	caacgttgycatatgcacctcgccgcgaatlgacctgagcctctgcctctcttaactatccc	974							
Db	1118	CTTGCCCAAGGGCGCTTTCACACGCTTACCACATCTTCCCTCGCTCGTACACAGCTTCAACAG	1177							
QY	975	cgaagctctcccaagtgcaccaagtccagcccgagcaaaagtctacagggggacatglatccc	1034							
Db	1178	CGTCCGCCAACCCCGCGTCTACTGCTGTGTCAGCGAAGCACCCACCGGAGCCTTGCCCG	1237							
QY	1035	caaccggggcgctgggc	1050							
Db	1238	CCTCCGCGGGGCGCTGC	1253							

RESULT 10
 US-08-687-289A-1
 : Sequence 1, Application US/08687289A
 : Patent No. 5981195
 : GENERAL INFORMATION:
 : APPLICANT: Fuller, Forrest H.
 : APPLICANT: Krapcho, Karen J.
 : APPLICANT: Hamerlund, Lance G.
 : TITLE OF INVENTION: CHIMERIC RECEPTORS AND METHODS FOR
 : TITLE OF INVENTION: IDENTIFYING COMPOUNDS ACTIVE AT
 : TITLE OF INVENTION: METABOTROPIC GLUTAMATE RECEPTORS AND
 : TITLE OF INVENTION: THE USE OF SUCH COMPOUNDS IN THE
 : TITLE OF INVENTION: TREATMENT OF NEUROLOGICAL DISORDERS
 : NUMBER OF SEQUENCES: 8
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Lyon & Lyon
 : STREET: 633 West Fifth Street
 : STREET: Suite 4700
 : CITY: Los Angeles
 : STATE: California
 : COUNTRY: U.S. A.
 : ZIP: 90071-2066
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 : MEDIUM TYPE: storage
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: IBM P.C. DOS 5.0
 : SOFTWARE: FastSED for Windows 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/687,289A
 : FILING DATE: July 25, 1996
 : CLASSIFICATION: 433
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 60/001,526
 : FILING DATE: July 26, 1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Warburg, Richard J.
 : REGISTRATION NUMBER: 32,327
 : REFERENCE/DOCKET NUMBER: 220/004
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (213) 489-1600

```

: TELEFAX: 7 (213) 955-0440
:
: TELEX: 67-3510
:
: INFORMATION FOR SEQ ID NO:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 3384 base pairs
:
: TYPE: nucleic acid
:
: STRANDEDNESS: single
:
: TOPOLOGY: linear
:
US-08-687-289A-1

```

Query Match	2.2%	Score 40.8	DB 2	Length 3384
Best Local Similarity	46.2%	Pred. No. 1.1		
Matches 135	Conservative 0	Mismatches 157	Indels 0	Gaps 0

QY	329	aagagacccgcgcggggaaccaggatcttcttccttttgaggaaacctgggccctctcgc	388
Db	2091	ACCCCAGTGTGTTGGGGCTAACCTCAGCTCGCTGCTGTTTTTCTCTGCACCTTCATGC	2150
QY	389	cgttgttgcctgttgtgtgaagcccgactctccaactgtgacctctgcgcgtctcct	448
Db	2151	AGATTGCATCTGTGTGATGTGGCTCACACGCCGCCCTCAAGCTACCGCAACGAGG	2210
QY	449	ttagggctcgtctgcgcacatctgtctctctgtctgagcgactaagtctttgscctcaact	508
Db	2211	ACCTGGAGATGAGATCATCTTTATACAGTGCCACAGGGCTCCCTCATGGGCCCTGGGCT	2270
QY	509	tccttgaccagaagaacaacacgagccccgggagcttgagtgacttccaacttbgcctctgcgc	568
Db	2271	TCCGTAGTCGGCTACACTGGCTGCTGGTGCCATCGCTCTTCTTTGGCTTCGAAGTCC	2330
QY	569	tgaacctggttagaggtcatcatcataacagagtgtgactgatatacaccctggt	620
Db	2331	GGAAGCTGCCGAGAACTTCATGAAGCCCAAGTCTATCACTTCCTGACCATGCT	2382

RESULT 11
 US-08-485-588-3
 Sequence 3, Application US/08485588
 Patent No. 5688938
 GENERAL INFORMATION:
 APPLICANT: Edward M. Brown
 APPLICANT: Steven C. Hebert
 APPLICANT: Forrest H. Fuller
 APPLICANT: James E. Garrett, Jr.
 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 TITLE OF INVENTION: MOLECULES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: First Interstate World Center
 STREET: Suite 4700
 STREET: 633 West Fifth Street
 City: Los Angeles
 STATE: California
 COUNTRY: USA
 Zip: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: FASTSEQ
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/485,588
 FILING DATE: 7 June, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 PRIOR APPLICATION DATA: Including application
 PRIOR APPLICATION DATA: described below: 9
 APPLICATION NUMBER: 08/353,784
 FILING DATE: 9 December, 1994
 APPLICATION NUMBER: PCT/US/94/12117
 FILING DATE: 21 October, 1994


```

: APPLICATION NUMBER: U.S. 08/292,827
: FILING DATE: 23 August, 1994
: APPLICATION NUMBER: U.S. 08/141,248
: FILING DATE: 22 October, 1993
: APPLICATION NUMBER: U.S. 08/009,389
: FILING DATE: 23 February, 1993
: APPLICATION NUMBER: U.S. 08/017,127
: FILING DATE: 12 February, 1993
: APPLICATION NUMBER: U.S. 07/934,161
: FILING DATE: 21 August, 1992
: APPLICATION NUMBER: U.S. 07/834,044
: FILING DATE: 11 February, 1992
: APPLICATION NUMBER: U.S. 07/749,451
: FILING DATE: 23 August, 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Heber, Sheldon O.
: REGISTRATION NUMBER: 38,179
: REFERENCE/DOCKET NUMBER: 213/005
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELETYPE: 67-3510
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3809 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 373..3606
: OTHER INFORMATION:
:
: US-08-485-588-3

```

```

Query Match          2.2%: Score 40.8; DB 1; Length 3809;
Best Local Similarity 46.2%; Pred. No.1.2; Indels 0; Gaps 0;
Matches 135; Conservative 0; Mismatches 157;
QY 329 aacgagacgtctgtggaaccaggtatctctcttctgggagacccttgagcctctctgac 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2516 ACCGCAAGTGTGGGGCTCAACCTGCAGTTCTGCTGTTCTCTGCGACCTTCATGCG 2575
QY 389 tcgtgttgcctgtgtgtgaagccgaactctccaccctgtgctctcgagcgtctctct 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2576 AGATTGTCATCTGTGTATCTGCTCTACACCGCGCCGCCCTCAAGCTACCGCAACGAG 2635
QY 449 ttggggttcgttcgcacatctgctctctctgtctgagcggtcgaagctcttgccctcaact 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2636 AGCTGGAGGATGAGATATCTTTCATCAGTCGTCACGAGGCGCTCCCTCATGCGCTTGAGCT 2695
QY 509 tccctgcccgaagaacacgagcccgagggtctggtgatcttcaactgtgctctgctgc 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2696 TCCTGATCGGCTACACCTGCGCTGCTGCGCATCTGCTTCTTTGCGCTCAAGTCCG 2755
QY 569 tgaccctgtgtagaggtcattcaatacagagtggtcgatcattcaacctgtgt 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2756 GGAAGCTGCCGAGAACTTCATGAGCAAGCTTCATCACCCTTCAGATGAGTCT 2807

```

```

RESULT 12
US-08-484-565-3
: Sequence 3, Application US/08484565
: Patent No. 5763569
: GENERAL INFORMATION:
: APPLICANT: Edward M. Brown
: APPLICANT: Steven C. Hebert
: APPLICANT: James E. Garrett, Jr.
: TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
:

```

```

: ADDRESSEE: Lyon & Lyon
: STREET: First Interstate World Center
: STREET: Suite 4700
: STREET: 633 West Fifth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: USA
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: FASTSEQ
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,565
: FILING DATE: 7 June, 1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: PRIOR APPLICATION DATA: including application
: PRIOR APPLICATION DATA: described below: 9
: APPLICATION NUMBER: 08/353,784
: FILING DATE: 9 December, 1994
: APPLICATION NUMBER: PCT/US/94/12117
: FILING DATE: 21 October, 1994
: APPLICATION NUMBER: U.S. 08/292,827
: FILING DATE: 23 August, 1994
: APPLICATION NUMBER: U.S. 08/141,248
: FILING DATE: 22 October, 1993
: APPLICATION NUMBER: U.S. 08/009,389
: FILING DATE: 23 February, 1993
: APPLICATION NUMBER: U.S. 08/017,127
: FILING DATE: 12 February, 1993
: APPLICATION NUMBER: U.S. 07/934,161
: FILING DATE: 21 August, 1992
: APPLICATION NUMBER: U.S. 07/834,044
: FILING DATE: 11 February, 1992
: APPLICATION NUMBER: U.S. 07/749,451
: FILING DATE: 23 August, 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Heber, Sheldon O.
: REGISTRATION NUMBER: 38,179
: REFERENCE/DOCKET NUMBER: 213/006
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELETYPE: 67-3510
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3809 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 373..3606
: OTHER INFORMATION:
:
: US-08-484-565-3

```

```

Query Match          2.2%: Score 40.8; DB 1; Length 3809;
Best Local Similarity 46.2%; Pred. No.1.2; Indels 0; Gaps 0;
Matches 135; Conservative 0; Mismatches 157;
QY 329 aacgagacgtctgtggaaccaggtatctctcttctgggagacccttgagcctctctgac 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2516 ACCGCAAGTGTGGGGCTCAACCTGCAGTTCTGCTGTTCTCTGCGACCTTCATGCG 2575
QY 389 tcgtgttgcctgtgtgtgaagccgaactctccaccctgtgctctcgagcgtctctct 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2576 AGATTGTCATCTGTGTATCTGCTCTACACCGCGCCGCCCTCAAGCTACCGCAACGAG 2635
QY 449 ttggggttcgttcgcacatctgctctctctgtctgagcggtcgaagctcttgccctcaact 508

```

Db 2636 AGCTGAGAGATGAGATCATCTTTCATCAGCTCCACGAGGCTCCCTCATGTGCCCCGCGCT 2695
Qy 509 tccctggcccggaagaacacgagcccgggctgggtgagtccttcacgtgtgctctgtcgcgc 568
Db 2696 TCCTGATGCGCTACACCTGCTGCTGCTGCCATCTGCTTCTTCTTCCCTTCAAGTCCC 2755
Qy 569 tgacctgtgtagaggtcalcatcaatacagagtgtgctgatacacccttgg 620
Db 2756 GGAAGCTCCGAGAACTTCATGAAAGCAAGTTTCATCACCCTTCAAGATGCT 2807

RESULT 13
US-08-480-751-3
Sequence 3, Application US/08480751
Patent No. 5858684
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,751
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-480-751-3

Query Match 2.2%, Score 40.8; DB 2; Length 3809;
Best Local Similarity 46.2%; Pred. No. 1.2;
Matches 135; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

Qy 329 aacgagacctgtgggagccaggtatcttcctctgtgggaacctgtgctctctgc 388
Db 2516 ACCGCACTGGTGGGGGCTCAACCTGCAGTTCCTCTGTTTTCCTGCACTTCATGC 2575
Qy 389 tcgtgttgacctgtgtgtgtaagccgacctctccacctgtgacctctgsgcttccct 448
Db 2576 AGATTGATCTGTGTGATCTGGCTCTACACGGGCCCCCTCAAGCTACCGCAACGAG 2635
Qy 449 ttgggtctctgtcgacatctctctctgtctgtgcgctcaagctcttgccctcaact 508
Db 2636 AGCTGAGAGATGAGATCATCTTTCATCAGCTCCACGAGGCTCCCTCATGTGCCCCGCGCT 2695
Qy 509 tccctggcccggaagaacacgagcccgggctgggtgagtccttcacgtgtgctctgtcgcgc 568
Db 2696 TCCTGATGCGCTACACCTGCTGCTGCTGCCATCTGCTTCTTCTTCCCTTCAAGTCCC 2755
Qy 569 tgacctgtgtagaggtcalcatcaatacagagtgtgctgatacacccttgg 620
Db 2756 GGAAGCTCCGAGAACTTCATGAAAGCAAGTTTCATCACCCTTCAAGATGCT 2807

RESULT 14
US-08-943-986-3
Sequence 3, Application US/08943986
Patent No. 5962314
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,986
FILING DATE: 03-OCT-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/484,565
FILING DATE: 7-June-1995
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117

```

? FILING DATE: 21 October, 1994
? APPLICATION NUMBER: U.S. 08/292,827
? FILING DATE: 23 August, 1994
? APPLICATION NUMBER: U.S. 08/141,248
? FILING DATE: 22 October, 1993
? APPLICATION NUMBER: U.S. 08/009,389
? FILING DATE: 23 February, 1993
? APPLICATION NUMBER: U.S. 08/017,127
? FILING DATE: 12 February, 1993
? APPLICATION NUMBER: U.S. 07/934,161
? FILING DATE: 21 August, 1992
? APPLICATION NUMBER: U.S. 07/834,044
? FILING DATE: 11 February, 1992
? APPLICATION NUMBER: U.S. 07/749,451
? FILING DATE: 23 August, 1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Heber, Sheldon O.
? REGISTRATION NUMBER: 38,179
? REFERENCE/DOCKET NUMBER: 213/006
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (213) 489-1600
? TELEFAX: (213) 955-0440
? TELEX: 67-3510
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 3809 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA to mRNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 373..3606
? OTHER INFORMATION:
US-08-943-986-3

```

Query Match 2.2%; Score 40.8; DB 2; Length 3809;

Best Local Similarity 46.2%; Pred. No. 1.2; Mismatches 157; Indels 0; Gaps 0;

```

Matches 135; Conservative 0;
QY 329 aacgagacgcctgctggagaccagatctctctctctgaggaccctctctcgc 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2516 ACCGCAAGTGTGGGGCTCAACCTGCAGTTCTCTGTTTCTTGACACTTCAATGC 2575
QY 389 tcgtgttgcctgtgtgtgaagcccgactctccacctgtgcctcgcgcctctc 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2576 AGATTGTCATCTGTGATCTGCTTACACCGCGCCCTCAAGCTACCGCAACGAG 2635
QY 449 ttgggggttcgttcgcacatctgtctctctctctctctctctctcgcctc 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2636 AGCTGGAGATGAGATCATCTTCATCACTGTCACGAGGGCTCCCTCATGCGGCT 2695
QY 509 tccctggccggaagaacacagggcccgaggcgtggtgactcttcacgtgtgcctc 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2696 TCCTGATCGGCTACACTGCTGCTGCTGCATCTCTTCTTTTCCCTCAAGTCCC 2755
QY 569 tgaccctgtgtagtgcatacaataacagagtggtcgtacatcacctcgt 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2756 GGAAGCTGCCGAGAACTTCATGAAAGCAAGTTCAATCACCCTTACGATGCT 2807

```

RESULT 15

US-08-353-784-3
Sequence 3, Application US/08353784
Patent No. 6011068

GENERAL INFORMATION:

APPLICANT: Edward F. Nemeth, Edward M.
APPLICANT: Brown, Steven C. Hebert,
APPLICANT: Bradford C. Van Wageningen, Manuel
APPLICANT: F. Balandrin, Forrest H. Fuller,
APPLICANT: Eric G. Delmar, and Scott T. Moe
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

```

? TITLE OF INVENTION: MOLECULES
? NUMBER OF SEQUENCES: 20
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Lyon & Lyon
? STREET: First Interstate World Center
? STREET: Suite 4700
? STREET: 633 West Fifth Street
? CITY: Los Angeles
? STATE: California
? COUNTRY: USA
? ZIP: 90071
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: FASTSEQ
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/353,784
? FILING DATE: 9 December, 1994
? CLASSIFICATION: 514
? PRIOR APPLICATION DATA:
? PRIOR APPLICATION DATA: including application
? PRIOR APPLICATION DATA: described below: 8
? APPLICATION NUMBER: PCT/US/94/12117
? FILING DATE: 21 October, 1994
? APPLICATION NUMBER: U.S. 08/292,827
? FILING DATE: 23 August, 1994
? APPLICATION NUMBER: U.S. 08/141,248
? FILING DATE: 22 October, 1993
? APPLICATION NUMBER: U.S. 08/009,389
? FILING DATE: 23 February, 1993
? APPLICATION NUMBER: U.S. 08/017,127
? FILING DATE: 12 February, 1993
? APPLICATION NUMBER: U.S. 07/934,161
? FILING DATE: 21 August, 1992
? APPLICATION NUMBER: U.S. 07/834,044
? FILING DATE: 11 February, 1992
? APPLICATION NUMBER: U.S. 07/749,451
? FILING DATE: 23 August, 1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Heber, Sheldon O.
? REGISTRATION NUMBER: 38,179
? REFERENCE/DOCKET NUMBER: 209/069
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (213) 489-1600
? TELEFAX: (213) 955-0440
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 3809 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA to mRNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 373..3606
? OTHER INFORMATION:
US-08-353-784-3

```

Query Match 2.2%; Score 40.8; DB 3; Length 3809;

Best Local Similarity 46.2%; Pred. No. 1.2; Mismatches 157; Indels 0; Gaps 0;

```

Matches 135; Conservative 0;
QY 329 aacgagacgcctgctggagaccagatctctctctctgaggaccctctctcgc 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2516 ACCGCAAGTGTGGGGCTCAACCTGCAGTTCTCTGTTTCTTGACACTTCAATGC 2575
QY 389 tcgtgttgcctgtgtgtgaagcccgactctccacctgtgcctcgcgcctctc 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2576 AGATTGTCATCTGTGATCTGCTTACACCGCGCCCTCAAGCTACCGCAACGAG 2635

```

OY 449 ttgggtctctgtcgcacatctgctctctctgtctgagcgctcaagtccttgccctcaact 508
|||
|||
Db 2636 AGCTGAGAGATGAGATCATCTTTCATACGTCGACAGAGGCTCCCTCATGSCCTGGGCT 2695
|||
|||
OY 509 tccctggcccgaaagaacacagggcccggggctggtgatacttcaactgtgctctgtgc 568
|||
|||
Db 2696 TCCGTGATCGGCTACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2755
|||
|||
OY 569 tgacccttgtagaggtcaatcaatacaaggtggtgatcatcacccctggt 620
|||
|||
Db 2756 GGAAGCTGCCGAGACACTTCAATGAAGCCAGTTCATCACCCTCAGCATGCT 2807
|||
|||

Search completed: September 22, 2002, 17:49:33
Job time: 11381 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2002, 14:38:27 : Search time 3328.72 Seconds
(without alignments)
11435.438 Million cell updates/sec

Title: US-09-895-686-7
Perfect score: 1819
Sequence: 1 cgcctcgagccctaccacgc.....ctttatctactcttaaaaa 1819

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_ov:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: gb_vl:*
16: em_ba:*
17: em_fun:*
18: em_hum:*
19: em_in:*
20: em_mu:*
21: em_mu:*
22: em_mu:*
23: em_mu:*
24: em_mu:*
25: em_mu:*
26: em_mu:*
27: em_mu:*
28: em_mu:*
29: em_mu:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query Match Length DB ID Description

1	1797	98.8	1866	6	AX083423	AX083423 Sequence
2	1786.8	98.2	1824	6	AF207989	AF207989 Homo sapi
3	1785	98.1	1808	6	AX136163	AX136163 Sequence
4	1753	96.4	1824	9	BC016860	BC016860 Homo sapi
5	1630.4	89.6	1790	6	AX014744	AX014744 Sequence
6	1357.4	74.6	1428	6	HS4276102	AJ276102 Homo sapi
7	1324.4	72.8	1326	6	AX014742	AX014742 Sequence
8	1323	72.7	1323	6	AX083413	AX083413 Sequence
9	1088	59.8	192087	2	AC079325	AC079325 Homo sapi
10	1068	59.8	197283	2	AC055863	AC055863 Homo sapi
11	826.8	45.5	191360	2	AL663079	AL663079 Mus muscu
12	826.8	45.5	196975	2	AL669969	AL669969 Mus muscu
13	679	37.3	1662	2	BC004925	BC004925 Homo sapi
14	575.8	31.7	613	6	AX136494	AX136494 Sequence
15	378	20.8	1845	6	AK000249	AK000249 Homo sapi
16	331	18.2	400	11	G14600	G14600 SHGC-11389
17	280.6	15.4	333	6	AX335205	AX335205 Sequence
18	270.6	14.9	77713	9	HUAC004131	AC004131 Homo sapi
19	270.6	14.9	134640	9	AC027130	AC027130 Homo sapi
20	270.2	14.9	1212	6	AX015106	AX015106 Sequence
21	270.2	14.9	1327	9	HS4276101	AJ276101 Homo sapi
22	270.2	14.9	2870	9	AF181862	AF181862 Homo sapi
23	270.2	14.9	3852	6	AX015108	AX015108 Sequence
24	268.6	14.8	1402	9	AF202640	AF202640 Homo sapi
25	228.4	12.6	251	6	AX136641	AX136641 Sequence
26	224.4	12.3	2899	10	BC020004	BC020004 Mus muscu
27	223.4	12.3	127219	2	AC023806	AC023806 Mus muscu
28	222.8	12.2	4488	10	AF378831	AF378831 Mus muscu
29	202.4	11.1	161577	9	AC007688	AC007688 Homo sapi
30	202.4	11.1	204272	2	AC007215	AC007215 Homo sapi
31	201.4	11.1	1619	6	AX078375	AX078375 Sequence
32	201.4	11.1	2302	9	AF095448	AF095448 Homo sapi
33	201.4	11.1	2305	9	BC003665	BC003665 Homo sapi
34	201.4	11.1	2446	9	AK001761	AK001761 Homo sapi
35	201.4	11.1	3371	6	AX188348	AX188348 Sequence
36	191.6	10.5	2484	6	AR129188	AR129188 Sequence
37	158.6	8.7	604	6	AX014842	AX014842 Sequence
38	144.4	7.9	1324	10	AF218809	AF218809 Mus muscu
39	140.6	7.7	519	6	AX079652	AX079652 Sequence
40	134.2	7.4	972	6	AX077889	AX077889 Sequence
41	132.6	7.3	1038	9	AF209923	AF209923 Homo sapi
42	132.6	7.3	1089	6	AX207667	AX207667 Sequence
43	132.6	7.3	161577	9	AC007688	AC007688 Homo sapi
44	117.2	6.4	2145	10	AB030197	AB030197 Mus muscu
45	99.4	5.5	47857	2	AC110297	AC110297 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS AX083423 1866 bp DNA
DEFINITION Sequence 115 from Patent WO0112660.
ACCESSION AX083423
VERSION AX083423.1 GI:13185264
KEYWORDS
SOURCE
ORGANISM human.

REFERENCE
AUTHORS Kato,S. and Kimura,T.
TITLE 1 (bases 1 to 1866)
JOURNAL Patent: WO 0112660-A 115 22-FEB-2001;
SAGAMI CHEMICAL RESEARCH CENTER (JP) ; Protegene Inc. (JP)
FEATURES
source location/Qualifiers
1..1866
/organism="Homo sapiens"
/db_xref="taxon:9606"
142..1467
/note="unnamed protein product"

```
/codon_start=1
/protein_id="CAC33306.1"
/db_xref="GI:13185265"
/translation="MAIHKALVMCLGIPLEFPVGNAGOHVPPGSGOGLNPLYYNICD
RSGAMGIVLEAVAGAGIVTFEVLTITLVASLPVQDTRKSLSTQVFLITGLGFLFC
LVACVYKPDFSTCASRRLFGLVLAFCESCLAHAFALFKRNHPRGWIFVLA
LLLTJLEVIINTMLITTLVRSGEGGPODSNAGNAVSPCAIAMDVFMALITVWL
LLLGAFGAMPALCGRYKRWKRKHGFVLTATSVAIWVWIVMYTGNQHNPSPTMD
DPTLALALANMAAFVLEFYIPEVSOVTKSPQSDQDMYPTRGVGETILKEOKGO
SMFEVKAFSEMPVAAKRPVSPYSGNGLSLSVOPTMALMHVYPSGAYDILLP
RATANSQVMSAISTILRAEDMYSASQSHQATTPPKDKNSVFRNPYVMD"
```

BASE COUNT 334 a 602 c 551 g 379 t

ORIGIN

Query Match 98.8%; Score 1797; DB 6; Length 1866;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1797; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

10 cccctaccagcggaaagtaacgaatcgctcagctcgaaggaaccaacagagctggc 69
|||||
70 CCCCTACCAGCGGAAAGTACGATCGGCTCAGCTGAGGAGCCAAACGAGACCTGGC 129
|||||
70 cttggagccaagatgagcaatccaaagccttggatgagctggagactgctctctc 129
|||||
130 CTGGAGCCACAGATGGCCATCCAAAGCCTTGGTGAATGCTGGGACTGCTCTCTTC 189
|||||
130 ctgtctccagagggcctggcccagaggaatgtcccaaccggctgcaagcgaagctaac 189
|||||
190 CTGTTCCAGGGGCTGGGCGCCAGGCGCATGTCACCCGGCTCAGCCAAAGGCGCTTCAAC 249
|||||
190 cccctgtactacaacatgctgagcagctctggagcgctggagcactgctccttgaagcgctg 249
|||||
250 CCCCTGTACTACAACTGTGTGACCGGCTGTGGGCGTGGGCATGCTCTGAGGCGCGTG 309
|||||
250 gctggagcgaggaatgtgcaacagcttctgctcacaatcactcgtgtggccagctccccc 309
|||||
310 GCTGGGGGCGGCAATGTGTACACACGTTTGTCTACATCATCTGTGGGCGCCAGCTCCCGC 369
|||||
310 ttgtgagagacacaaagaaacgagcctgtctggagacccagatctctctctcttgaagg 369
|||||
370 TTGTGTGAGGACACCAAGAAACGAGCGCTGCTGGGAGCCAGATATCTTCTTCTTGGGG 429
|||||
370 accctggagcctctctgctcgttctgctcgttggatgagacccgaacttctccacctgt 429
|||||
430 ACCCTGGGCGCTTCTGCTGCTGTGGTGGCTGTGGTGAAGCCGCACTTCTCCACCTGT 489
|||||
430 gectctggagcctctctcttggagctctgtctgacatctgctctcttcttctggcgct 489
|||||
490 GCTCTCTGGGCGCTTCTCTTGGGGTCTGTTCGCCATCTCTCTTGTCTGGGCGCT 549
|||||
490 cagctcttggcctcaaatctcttggcccggaagaaacagagcccgaggagtgagtgatc 549
|||||
550 CAGCTCTTGGCTTCACTTCTGCGCGGAGAACACAGGGGCCCGGGGCTGGGTGATTC 609
|||||
550 ttcactgtgctcgtcgtctgacccgttaaggttcatlcaatcaatagagtgagctgac 609
|||||
610 TTCACTGTGGCTGTGCTGTACCTGTAGAGTCAATCATATACAGATGGCTGATTC 669
|||||
610 atcaacctgtgtcggagcagtgagagcgccctcaaggcaacagcagcagcgctgag 669
|||||
670 ATCAACCTGTTGGGCGAGTGGGAGGCGGCGCTCAGGCGCAACAGCAGGCGAGGCTGG 729
|||||
670 gcgatggcctccctctgacatctgcaacaatgagacttgcagtgagcaatctatagctc 729
|||||
730 GCGTGTGGCTCCCTCTGTGCGCATGCGCAACATGAGCTTGTGATGAGCATATATAGCTTC 789
|||||
730 atgtctgtcgtcgtcgtgctcctctcttggagggcgcccgcccgctgtgtgagcgctacaag 789
|||||
790 ATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 849
|||||
790 cgtctgagtaagcattggagcttctgtctctctcaacagcgaactcgttgcataatg 849
|||||

Db 850 CGTGGCGTAAGCATGGGGTCTTGTGCTCTCTCAACAGCCACCTCCGTTGCCATATGC 909
850 gtgggtgagatcgctcatctacttaccgcaacaagacgaacaaatcccaactggagat 909
|||||
910 GTGGTGTGATCGTCATGTATACTTACGGCAACAAGACAGCAACAGTCCACCTGGGAT 969
|||||
910 gaccacagctgtgacatcgccctcgcgcgaatgacctggagccttctctcttcttaccg 969
|||||
970 GACCCACAGCTGGCCATCGCCCTCGCCGCAATGCTGTGGGCTGTGCTCTTCTTACGTC 1029
|||||
970 atccccgaggtctcccggtgaccaaagtcgaagcccaagcaagaagctcccgaggagatg 1029
|||||
1030 ATCCCGAGGTCTCCAGGTGACCAAGTCCAGCCCAAGCAAGCTACCAAGGGGACATG 1089
|||||
1030 taaccacaccgggagcggtgagctatgagaccaatccctgaagaagcagaaggttcaagatg 1089
|||||
1090 TACCCACACCGGGCGGTGGCTATGAGACCATCTGTGAAGAGCAAGAGGTCAACAGCATG 1149
|||||
1090 tctgtgagaaacagagccttccatgagatgagccggttgcagctaaagagcggtgtca 1149
|||||
1150 TTGCTGGAACAAGAGCGCTTTCATGATGAGCGGTGACGCTTAAGAGCGCGGTGCA 1209
|||||
1150 ccatcaagcggtgtacaatggagcagctgtctgacagatgtgtacagcccaatgagatggc 1209
|||||
1210 CCATACAGCGGGTCAATGGGACACTGTGACAGTGTGTACAGCCCACTGAGATGGCC 1269
|||||
1210 ctgagtcaaaagtctcgctccgaagagctttagacatcctcccaagcgccacggc 1269
|||||
1270 CTGATGCAAAAGTTCCGTCCGAAGAGACTTACACATCATCTCTCCACGGGCGCACCGCC 1329
|||||
1270 aacagcaggtgagatgagcagtgccaactcgacccctggcggtgaaagacatgtactcgcc 1329
|||||
1330 AACAGCCAGGTGATGGGAGTGCACACTGCACCTCGGGGCTGAAGACATGACTCTCGGCC 1389
|||||
1330 cagaagcaacagagcgagcgaacacgcgcgaagaagcgcaagaactctcaggtctttagaac 1389
|||||
1390 CAGAGCCACCAAGGGGCGCACCGCCCAAAAGACGGAAGACTCTCAAGTCTTTAAGAAC 1449
|||||
1390 cccctagctgtgagactgagtcagtcggtggagagagagcggtctgagatcttggagagggc 1449
|||||
1450 CCTACGTGTGGAGTGAAGTACGGGTGCGAGAGAGAGCGGTGGATTGGGAGGGGCC 1509
|||||
1450 ctgagagactggcccgggcaagaggaacttcagagctctctccctctggcagggccag 1509
|||||
1510 CTGAGAGACTGTGCCCCGGGCAAGGAGACTTCAGAGCTCTCTCCCTGCGCAGGCGCAG 1569
|||||
1510 caaatgtgcccagaatgtgaaagagcctccctctctgcagatgttgggtgtgtcat 1569
|||||
1570 CAACATGTGCCCCAGATGTGAAGGGCTCCCTCTGTGCCAGTGTGGGTGGGTGTAT 1629
|||||
1570 ggtgtcccccacccactccctcagtgcttggagtgagagagcaaccccaagcctctgac 1629
|||||
1630 GGGTGTCCCCACCCACTCTCTAGTGTGTGGATCCAGAGGCAACCCAGCTCCTGCG 1689
|||||
1630 caagatcaactcggcgggtlcaacactccagcaaatagtgcttcgggtgtgtgagtgagca 1689
|||||
1690 CAGGATACACTCGGCGGTACACTTCAGCCAAATAGTGTCTCGGGGTGGGTGCTGGGCA 1749
|||||
1690 ggcgtatgttctctctgtgagatctccgtcaactcaagaagatctcccaagcgctcaagcct 1749
|||||
1750 GCGGCTATGTTTCTCTGTGAGATTCTCAACCTCAAGAGATTCTCCAGGCGCTCAGGCT 1809
|||||
1750 ggaactgtactctctgtgaggaacaagagtgacctaatataatcttctgtcttact 1806
|||||
1810 GGATCTGTCTCTGTGTGAGAACAGAGGTGCTTAATAATACATTTCTGCTTTATT 1866
|||||

RESULT 2
AF207989 1824 bp mRNA linear PRI 09-MAY-2001
LOCUS AF207989
DEFINITION Homo sapiens orphan G-protein coupled receptor (GPR5C) mRNA,
complete cds.
ACCESSION AF207989

VERSION	AF207989.1	GI:8118031
KEYWORDS	human.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 1 to 1824) Brauner-Osborne,H., Jensen,A.A., Sheppard,P.O., Brodin,B., Krosgaard-Jaensen,P. and O'Hara,P.	
TITLE	Cloning and characterization of a human orphan family C G-protein coupled receptor GPRC5(1)	
JOURNAL	Biochim. Biophys. Acta 1518 (3), 237-248 (2001)	
MEDLINE	21210986	
PUBMED	11311935	
REFERENCE	2 (bases 1 to 1824) Brauner-Osborne,H., Sheppard,P.O. and O'Hara,P.J.	
AUTHORS	Direct Submission	
TITLE	Submitted (23-NOV-1999) Department of Medicinal Chemistry, The Royal Danish School of Pharmacy, 2 Universitetsparken, Copenhagen 2100, Denmark	
JOURNAL	Location/Qualifiers	
FEATURES	1..1824	
source	/organism="Homo sapiens"	
gene	/db_xref="taxon:9606"	
CDS	1..1824 /gene="GPRC5C" 83..1408 /gene="GPRC5C" /note="similar to RAIGL, GPRC5B and GPRC5D" /codon.start=1 /product="orphan G-protein coupled receptor" /protein_id="AAF2870.1" /db_xref="GI:8118032" /translation="MAIHKALVMGLPLFLPGAMAQHVPDSCSGNLNPLYNLCDSFGWAGYIVLDAVAGAGIVTTFVLIIIVASLPFQODIKRSLIGTQVFLITGLTGLFC LVFCAVCYDPDSTSCASRPLFGLVFAIFSCGLAEHVALNPLANKINGPBGWFTVA LLLTLVEIITLFWLITLTVRSGSEGGPGCNSSGAMVASPCALAMNDPMALITVWL LLLCAFLGAMPALCGRTKRRKRHVEYFLLTATSVAlNYWYIWTYITGNKQNSPTWD DPTLIALAANAAMFLLVYIPEVSQYTKSSPEGSYGDMTPTKGVETITLAEKQD SMFEAKFASMDPEVNAKRPVSPYSGYNQGLISVYPTDMALMHKVPSCAVDIILP RATNSOVMSANSSTLRAEDMYSASHQAAATPPKDGKNSQVFRNPYVMD"	
BASE COUNT	337 a	579 c 535 g 373 t
ORIGIN		
Query Match	98.2%	Score 1786.8; DB 9; Length 1824;
Best Local Similarity	99.6%	Pred. No. 0;
Matches 1802; Conservative	0;	Mismatches 7; Indels 1; Gaps 1;
QY	10	ccctaccagccgggaagtacagatcgcgctcaagctctgtgaaggagaccacaacagagcctggc 69
DB	11	ccctaccagccgggaagtacagatcgcgctcaagctctgtgaaggagaccacaacagagcctggc 70
QY	70	ctggagaccagatgagccatccacaagaacgctgtgtatgttcgtctgagatgcctctcttc 129
DB	71	ctggagaccagatgagccatccacaagaacgctgtgtatgttcgtctgagatgcctctcttc 130
QY	130	ctgtctccagagggcctgtggccagagagcatgtccaccgccgagctgacgaccaagagctcaac 189
DB	131	ctgtctccagagggcctgtggccagagagcatgtccaccgccgagctgacgaccaagagctcaac 190
QY	190	cccctgtactacaacctgtgtaacgcgctcgtggggcgctggggagatcgtctcttgaagccgtg 249
DB	191	cccctgtactacaacctgtgtaacgcgctcgtggggcgctggggagatcgtctcttgaagccgtg 250
QY	250	gctggggcgggagatgtgtaccacgcttctgtctacacatcattcctgtgtgccaagcctccc 309
DB	251	gctggggcgggagatgtgtaccacgcttctgtctacacatcattcctgtgtgccaagcctccc 310
QY	310	tttgtgcaggaacccaagaagagagcctgtctgtggagaccagagtattctctctcttgagg 369
DB	311	tttgtgcaggaacccaagaagagagcctgtctgtggagaccagagtattctctctcttgagg 370

[illegible]

Db	1451	CTGAGACTGGACCCCGGGCAAGGGAAGTCTCCAGGCTCCTCCCCCTGGCAGG-CCAG	1509
Qy	1510	caacatgtgcccccaabatytgaaagggcctccctctctccaggltgttggtygtgat	1569
Db	1510	CAACATGTGCCCCAGATGTGGAAGGCGCTCCTCTCTGCAGTGTTCGATTTGGTGTCAT	1569
Qy	1570	gggtgtcccccaccactctcagtgltttgttgagtgcgaagagacaaccagctcttcg	1629
Db	1570	GGGTGTCCCACCACCTCCTCAGTGTTCGTAAGTGCAGAGACCAACCCACCTCTGC	1629
Qy	1630	caggataacctcggcgtcacactccaagcaatatgtgtctcgaggtggtgtgcgagca	1689
Db	1630	CAGGATACACTTGGCGGTGCACACTCCAGCAATAAGTGTCTCGGGGTGTGGCTGGGCA	1689
Qy	1690	gcgcctatgttctctctygagattctctgcaacctcaagaagactlcccgagcgctcagcct	1749
Db	1690	GGCGCATGTTCCTCTCGAGATTTCCTGCNAACCTCAAGAGACTTCCAGGCGTCAGGCT	1749
Qy	1750	ggatcttgtctctctgttgagaacaagggtgccttaataatatlctgtcttatc	1809
Db	1750	GGATCTGTCTCTCTGTGTGAGAACAAGGTGCTTAATAATCATTTCTGCTTTATTAA	1809
Qy	1810	tcttaaanaa 1819	
Db	1810	AAAAAAAAA 1819	
RESULT	3		
LOCUS	AXJ36163	1808 bp	DNA linear PAT 30-MAY-2001
DEFINITION	Sequence	85 from Patent EPI067182.	
ACCESSION	AXJ36163		
VERSION	AXJ36163.1	GI:14272571	
KEYWORDS	SOURCE	human.	
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Ota,T., Isogai,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and Hayashi,K		
JOURNAL	Secretory protein or membrane protein		
FEATURES	Patent: EP 1067182-A 85 10-JAN-2001;		
source	Helix Research Institute (JP)		
CDS	Location/Qualifiers		
	1..1808		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	85..1410		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC39724.1"		
	/db_xref="GI:14272572"		
	/translation="MAIHKALVWLCLPLLPFGAMAOGHVPECCSGGLNPLYNLCD		
	/rsganglaeion="MAIHKALVWLCLPLLPFGAMAOGHVPECCSGGLNPLYNLCD		
	RSGANGIALEAVAGAGIVTFVLITLLVASLPEVDTKRSLGTQVEFLGTGLFCC		
	LVPACVVWPDFSTCASRRFLFEGFAIFCSCLAHYVALLNARKNGPRGWLFYVA		
	LLLTVEVINTENTMLIITLVRSGEVGGOSSGMVAASPCAIANNDFVALITYVM		
	LLLAIFGAPALICGRYKRMRKHGVFVILITATSVAILIMVVIWVTYGNOHNSPTMD		
	DPTLAIALANAAMAFVLFYVLPESVOTKSSPESDYSOGDMPTPGVGSEFTILKOKCO		
	SMFENKASMEDEPVAAARVPSPYSGYNQLITTSVYOETEMALMKHPSEGADVIDILP		
	RATANSQVMSANSITLRATEDMYSAOSHQAATPPDKNSQVFRNPYYWD"		
BASE COUNT	321 a	580 c	531 g 376 t
ORIGIN			
Query Match	98.1%:	Score 1785;	DB 6; Length 1808;
Best Local Similarity	99.9%:	Pred. No. 0;	
Matches 1796:	Conservative	0; Mismatches	0; Indels 1; Gaps 1;
Qy	10	ccctcaacgagccggaagtagtgagtcgagctcagctcgtgagggagcccaacagagctcgtgc	69
Db	13	CCCTCAACGAGCCGGAAGTAGTGAGTCGAGCTCAGCTCGTAGGAGGAGCCCAACAGAGCTTGGC	72

OY	70	cttgagagccagagatgagccatccaaagaacctggtatgtatgtcttgaggactgctctctc	129
Db	73	CTGGAGGCAAGAGTGGCCATTCACAAAGCCTTGATGTATGCTTGAGACTGCTCTCTTC	132
OY	130	ctgttcccaagggagccttgagcccaaggagccatgtctccaccggcgtctgcagcccaagcctccac	189
Db	133	CTGTTCACAGGGGCGCTGGGCCAAGGGGCATGTCCACCCGGGTGAGGCCAAGGCGCTCAAC	192
OY	190	ccccctgtactcaaacctgtgtgtaacgctctctggagcgttgaggcaatcgctcttgagacgtg	249
Db	193	CCCCGTACTAACACCTGTGTGACCGCTGTGGGGCTGGGGCATGCTCTCTGAAGGCGGTG	252
OY	250	gctgggagcgagcatgtgtcacacagctltgtgtcacaatcatcctgttgagccagcctccccc	309
Db	253	GCTGGGGGGGSCATTGTGCACACAGTTTGTGTCTCACCATCATCTGTGTGGCAGGCTCCCC	312
OY	310	cttgtgtcagagcaccaagaagacgagcgtctggtggagaccaggtatctctctcttgagg	369
Db	313	TTTGTGGCAGGACACCAAGAAACGAGCGCTGTGGGACCCAGGTATTTCTTCTTGTGGGG	372
OY	370	acctctggagccctctctgtccctcgtgtttgtgcctgtgtgtgaagccagactctccaccgt	429
Db	373	ACCCTGGGCGCTCTTCTCTCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	432
OY	430	gctctctcgagcgtctctctctctctctctgtgtgtctgtctgtcgcacatctctctgtctgagcgt	489
Db	433	GCTCTCGGCGCCTTCCTTGTGGGGTTCTGTTCGCCATCGTTCCTTCTGTCTGTGGCGCT	492
OY	480	caagctcttctgcctccaactctctctgtgcccagagaaacaacagggcccgagcgtgtgtatc	549
Db	483	CACGCTTTTGCCCTCAACTCTTCGTGGCCGGAAAMACACGGGCGCCGGCGGTGTGATC	552
OY	550	ctcaactgtgctctgtctgtctgtacacccgtgtagaagtcataatcaatacaagatgtctgtatc	609
Db	553	TTCACTGTGTGCTGTGCTGTGACCCGTGTGATGAGTGTATCATCAATATACAGATGGCTGTAC	612
OY	610	atcaaccctgtgtctggggcagltgagccagagcgagcctcaaggcaacagcagcgagcgtgtg	669
Db	613	ATCACCCCTGTTCGGGGCAGTGGCGAGGGGGGCGCTTCAGGGCAACAGCGACGCAAGCTGTG	672
OY	670	gcgcgtgagcctcccccctgtgtcacagccaaatgagactttgtcatgtgaactctactagctc	729
Db	673	GCGCTGGGCTCTCCCTCTTGCCATTCGCCACATGTGACTTTGTATGTGCACTATCTACGTC	732
OY	730	atgtcgtctgtctgtgtgtgtcctctctgtggggccctgtgcccgcctctgtgtgcccgtacaa	789
Db	733	ATGCTGCTGCTGCTGTGGGTGCTTCCTGTGGGGCGCTGGCCGCTGTGTGTGGCGCTACAG	792
OY	790	cgctgtggttaagcaatgggtgtcttgtgtctctcaacagccacactcgtgtgtccatatgg	849
Db	793	CGCTGGGCTAAGCAATGGGGGCTTTGTGTCTCTTACACAGCCACCTCGTGTCTTCCATATGG	852
OY	850	gttgtgtgtgaatgtcagtaacttaactaacgcaacaagcagcacaaagcttcccaactggagat	909
Db	853	GTTGTGTGTGATCGTCAATGTATACTACGGCAACAACACAGCACAAAGTCCACACTGGGAT	912
OY	910	gaacccacgctgtgccaatcgccctcgtccccaatgtcgttgagcctgtgcctctctactagtc	969
Db	913	GACCCCAAGGCTGTGCAATCGCCCTTCGCGCCCAATGCGCTGGGCTTCGTCTCTTCTACACTC	972
OY	970	atcccgagaggtctcccaaggtgtacccaagttcacagccagacaaagcttcaacaaggaggacgt	1029
Db	973	ATCCCCGAGTCTTCCAGGTGACCAAGTCCAGCCCAAGACCAAAAGCTACCAAGGGGACATG	1032
OY	1030	taccaccaaccggggcgtgtgggctctatgtgagacatctctaaagcagacagaggttcaaggacatg	1089
Db	1033	TACCCACACCCGGGCGTGGGCTATGAGACATCTCTAAAGAGACAGAAAGGTCTCAGAGCATG	1092
OY	1090	cttgtgtgagaacaagcctcttccatgtgatgtgacggtgtgtcagctcaagagccggtgtctca	1149
Db	1093	TTCTGTGAGAAACAAGGCTTTTCTCATATGATGATGACCGGTGTGACGTAAAGGCGCGGTGTCA	1152

Contact: (Dickson, Mark) mdc@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAK Plate: 22 Row: h Column: 1.

FEATURES

Location/Qualifiers

```
/organism="Homo sapiens"
/db xref="LocusID:55890"
```

```
/usr_x86_64 = laxon:9606
```

```
/tissue_type="Pancreas, epitheli
```

```
/lab host="DH10B"
```

```

/home/Vector: pcmv-3P0R16
34 1340

```

```
/codon_start=1
```

member C"

/protein_

```
/translation="
```

12345678910111213141516171819202122232425262728293031323334353637383940414243444546474849505152535455565758596061626364656667686970717273747576777879808182838485868788899091929394959697989910010110210310410510610710810911011111211311411511611711811912012112212312412512612712812913013113213313413513613713813914014114214314414514614714814915015115215315415515615715815916016116216316416516616716816917017117217317417517617717817918018118218318418518618718818919019119219319419519619719819920020120220320420520620720820921021121221321421521621721821922022122222322422522622722822923023123223323423523623723823924024124224324424524624724824925025125225325425525625725825926026126226326426526626726826927027127227327427527627727827928028128228328428528628728828929029129229329429529629729829930030130230330430530630730830931031131231331431531631731831932032132232332432532632732832933033133233333433533633733833934034134234334434534634734834935035135235335435535635735835936036136236336436536636736836937037137237337437537637737837938038138238338438538638738838939039139239339439539639739839940040140240340440540640740840941041141241341441541641741841942042142242342442542642742842943043143243343443543643743843944044144244344444544644744844945045145245345445545645745845946046146246346446546646746846947047147247347447547647747847948048148248348448548648748848949049149249349449549649749849950050150250350450550650750850951051151251351451551651751851952052152252352452552652752852953053153253353453553653753853954054154254354454554654754854955055155255355455555655755855956056156256356456556656756856957057157257357457557657757857958058158258358458558658758858959059159259359459559659759859960060160260360460560660760860961061161261361461561661761861962062162262362462562662762862963063163263363463563663763863964064164264364464564664764864965065165265365465565665765865966066166266366466566666766866967067167267367467567667767867968068168268368468568668768868969069169269369469569669769869970070170270370470570670770870971071171271371471571671771871972072172272372472572672772872973073173273373473573673773873974074174274374474574674774874975075175275375475575675775875976076176276376476576676776876977077177277377477577677777877978078178278378478578678778878979079179279379479579679779879980080180280380480580680780880981081181281381481581681781881982082182282382482582682782882983083183283383483583683783883984084184284384484584684784884985085185285385485585685785885986086186286386486586686786886987087187287387487587687787887988088188288388488588688788888989089189289389489589689789889990090190290390490590690790890991091191291391491591691791891992092192292392492592692792892993093193293393493593693793893994094194294394494594694794894995095195295395495595695795895996096196296396496596696796896997097197297397497597697797897998098198298398498598698798898999099199299399499599699799899910001001100210031004100510061007100810091010101110121013101410151016101710181019102010211022102310241025102610271028102910301031103210331034103510361037103810391040104110421043104410451046104710481049105010511052105310541055105610571058105910601061106210631064106510661067106810691070107110721073107410751076107710781079108010811082108310841085108610871088108910901091109210931094109510961097109810991100110111021103110411051106110711081109111011111112111311141115111611171118111911201121112211231124112511261127112811291130113111321133113411351136113711381139114011411142114311441145114611471148114911501151115211531154115511561157115811591160116111621163116411651166116711681169117011711172117311741175117611771178117911801181118211831184118511861187118811891190119111921193119411951196119711981199120012011202120312041205120612071208120912101211121212131214121512161217121812191220122112221223122412251226122712281229123012311232123312341235123612371238123912401241124212431244124512461247124812491250125112521253125412551256125712581259126012611262126312641265126612671268126912701271127212731274127512761277127812791280128112821283128412851286128712881289129012911292129312941295129612971298129913001

LLLTVEVILNTEWLLITLVKGS

DP7LAIALANAWAEVLEYVIPLEVSQVTKSSPEQSYQGDMPTRGVGYETILKEQKGC

RATANSQVMSGANSTLRADMYSAOSHQAATPPKDGKNSOVERNPYVWD

Query Match 96.4%; Score 1753; DB 9; Length 1824;

```
Matches 1756; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

59 cagaacctgacctgaggaaccagatgccaatccacaaaccttgatgatattacctgagac

1 CAGAGCTGCTGGAGCCAGGATGCTATCCACAAGCTTGGTGAATGTCCTGGGAC 60

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

QY	1679	gtgagctggagcagcgccctatgcttctctcggagatctccgaacctaaagagactcccaag	1738		
Db	1621	gtggctgggagccgcctatgattttctggagattctgcCAACCTCAGAGACTTCCAGG	1680		
QY	1739	cgcctagcgcctggatctctgcctctctgtgaggaacaagggtgcctaaataacattctc	1798		
Db	1681	cgcttagcccttgatcttgcctcctctctgtgaggaacCAAGGGTGCCTAATAATACATTTC	1740		
QY	1799	gctttattactcttaaaaaa	1819		
Db	1741	gctttatttAAAAAAAAAAAAAA	1761		
RESULT	5				
AX014744	AX014744	1790 bp	DNA	linear	PAT 07-SEP-2000
LOCUS	Sequence 3	from Patent WO9953054,			
DEFINITION	AX014744				
ACCESSION	AX014744				
VERSION	AX014744.1	GI:10041010			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1790)				
AUTHORS	Michaelovich,D., Hill,J., Medhurst,A. and Pangalos,M.				
TITLE	AX014744				
JOURNAL	Patent: WO 9953054-A 3 21-OCT-1999;				
FEATURES	SMITHLINE BEECHAM PLC (GB)				
source	1..1790				
BASE COUNT	324 a	560 c	522 g	379 t	5 others
ORIGIN	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
Query Match	89.6%;	Score 1630.4;	DB 6;	Length 1790;	
Best Local Similarity	97.7%;	Pred. No. 4.4e-297;			
Matches 1758;	Conservative	0;	Mismatches 26;	Indels 15;	Gaps 10;
QY	13	tcacagcgcggaagctacagctcgagctcagccttgaaaggacccaacagagcctgagctg	72		
Db	1	TCACAGCGCGGAAGCTAGAGATCGCGCTCAGCGCTGAGAGACCACAGAGCCTGGCGCTG	60		
QY	73	ggaagcagagatgagcatcacacaagcctgtgtatgtctgtgagactgctctctctctg	132		
Db	61	GGAGGCAGATTTGGCATTCACAAAGCCTGGTGTATGTGCTGGGACCTGCTCTCTCTG	120		
QY	133	ttccagagagagccttgagcccaaggagcattgtccaccggagctcgagagccaaggcctcaaccc	192		
Db	121	TTCCAGAGGAGCCTGGG--CAAGGCAGATGTCCACCCGAGCTGCAG--CAAGGCTCAACCCC	177		
QY	193	ctgtactacaacacctgtgtgaacgctctcggagcgcttgaggacatgctcttgaggagcgctgtgc	252		
Db	178	CTGTACTACAACTGTGTGAGACCGCTTGAGGCGGATCGTCGAGAGCGCGTGGCT	237		
QY	253	ggagcgggcatgtgtgacccaaggttgtgtcacaatcatctctgtgtgcagagcctccctct	312		
Db	238	GGGGGGGGCATTTGTCACACACTTGTGTCTCAACATCTCTCGGTGGCCAGCCCTCCCTTT	297		
QY	313	gtgcagagacacaagaacagagagcctctgaggagccagagatcttctccctcttgaggac	372		
Db	298	GTGCAGAGACACCAAGAAAGAGGCTCTCTGGGAGCCAGATATTCTTCTCTGGGTACC	357		
QY	373	ctggagcctctctgctcctgtgttctgcctgtgtgtgagccagagcctctcaacatgtgc	432		
Db	358	CTGGGCGCTTCTGTCGCTCGTGTGTGTGCTTGTGCTGATGTGAAGCCGACATT--TCTCTGTGCC	414		
QY	433	ctctggagctctctctcttcttgagggtctgtgtgcacatctgctctctctgtctggagctac	492		
Db	415	TCTGTGGGCTTCTCTTGTGGGTCTGTGTGCGCATCTGCTTCTGTCTGTGGGCGCTCAG	474		

LOCUS	AX014744	1790 bp	DNA	linear	PAT 07-SEP-2000
DEFINITION	Sequence 3 from Patent WO9593054.				
ACCESSION	AX014744				
VERSION	AX014744.1	GI:10041010			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 1790)				
TITLE	Michaelovich,D., Hill,J., Medhurst,A. and Pangalos,M.				
JOURNAL	AX04.4-protein-coupled receptor Patent: WO 9953054-A.3 21-OCT-1999; SMITHKLINE BEECHAM PLC (GB)				
FEATURES	Location/Qualifiers				
Source	1..1790				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	324 a 560 c 522 g 379 t			5 others	
ORIGIN					
Query Match	89.6%; Score 1630.4; DB 6; Length 1790;				
Best Local Similarity	97.7%; Pred. No.4,4e-297;				
Matches 1758; Conservative	0; Mismatches 26; Indels 15; Gaps 10;				
QY	13	tcacagcagcgaagaagtacagctgagctcagccctcggagaagaccccaacagagcctgagcctg	72		
Db	1	TCACCAGCCGGAAAGTACGAGTCCGGCTCAGCTTGGAGGAGCCACAGACAGCTGGCCTG	60		
QY	73	ggaagcagaatgagccatccacaagaaccttgatgatgtaccttgagactgcctctctctctg	132		
Db	61	GGAGCCAGATTGGCATTTCCACAAAGCCTTGGTGTGATGTGCTGGGAGCTGCTCTTCTTG	120		
QY	133	ttcccaagagagccttgggcccagagccatgctcccaaccggcctgcagagccaagagcctcaaccc	192		
Db	121	TTCCCAAGGAGGCGCTGGG--CAAGGCAATGTCCACACCGCGTGCAG--CAAGGCTTCAACCCC	177		
QY	193	ctgtactcaaacctgtgtgaacgcctctggggcgttgaggacatctccttgagagcgtgagct	252		
Db	178	CTGTACTCAACACGCTGTGTGACCGCTCTGGGGCGTGGGGCATGCTCCGAGAGCGCGTGGCT	237		
QY	253	ggggcgggacattgcaccaagtttggtctaccaatcaatcctggtggcagcctccctt	312		
Db	238	GGGGCGGGCATTTGTACACACGCTTGTGTCTCAACATCATCTCTGGTGGCAAGCCCTCCCTTT	297		
QY	313	gtcagagacaccaaagaacggagcctctgggagccaggtatctcttcctctctggaggac	372		
Db	298	GTGCAAGACACCAAGAAACGAGGCTCTCTGGGAGCCAGGATATTCTTCTTGGTATACC	357		
QY	373	ctggagcctctctgctcgtgtgttcacctgtgtgtgaagccgagactctcaacatgtgac	432		
Db	358	CTGGGCGCTTCTGTGCGTCTGTGTGCTTGTGCTGTGTGTAAGCCGACATT--TCTCTGTGCC	414		
QY	433	ctctggcgtctcctctcttgggggtctggttcgcacatctgctctctctgtctggcagctaac	492		
Db	415	TCTGGGCGCTTCTCTTGTGGGTCTGTGTGCGCATTCGCTTCTTCTGTCTGGGCGCTCAG	474		

QY	493	gtcttgcctccaacttccttgcccggaagaaacacacgagcccccgggagcttgggtgatcttc	552
Db	475	gtcttttgcctccaacttccctgagcccggaagaaacacagggccccgggggctggtgatcttcc	534
QY	553	actgtgctctgtctgtcgaacccgtgtagaaggtcatalcataaacaagatgtgctgatacc	612
Db	535	actttgtgctctgtcgtcgtgacacccctgtgtagaagttatcatatcaatcagatgtcatc	594
QY	613	aaccttggcttcggggagcaatggcgagagcgccctcaaggagaaagaagcgagctggggcc	672
Db	595	accctggtttggggagcaactgtggcgaaggccggccctcaaggagaaagaagcgagctggggcc	654
QY	673	gttgacctcccccctgtgtgcacacgcacaatgatgacttgcatagtgcactatcagctatg	732
Db	655	gtggccttccccctgtggtgcacatcgcccaacatngaatcttntcatgsgactcatctacgtat	714
QY	733	ctgtctgtctgtgagtgatccttcctctgggggcccgtgcccgcgtgtgtgscgtacaagcgc	792
Db	715	ctgctgctgctgtggtgtgctcttctggtgggggcttgccctgctgtggtggccgacaaagcgc	774
QY	793	ttgggtagaagatagggtggtcttggctgcctccccaacaagcgacctcgtttgcatatgggtg	852
Db	775	ttggggtatagaaatgggggtcttggctgctntctacacacagccactcctggttccatattgggtg	834
QY	853	gttggtgatcgtcatgttactatcagcgacaacagcaacaaacagtcaccaacctggagatcac	912
Db	835	gtgtggatgtcgtcatgtatcttactttagggcacaacagacacaaatgctccacactgggatcac	894
QY	913	cccaagcttggccatcgccctcgccgccaatgcctgggcccctgcgtcctctcttcaagctalc	972
Db	895	cccaagcttggccatcgccctcgccgccaatgcctgggcccctgctgcttcttcttcaagctalc	954
QY	973	ccgaaggtctcccaaggtgacacaagttcccaagcacaagtaagtaacaggggagaaatgac	1032
Db	955	cccgaggtcttcccaaggtgacacaagttcccaagcacaagtaagtaacaggggagaaatgac	1010
QY	1033	cccaagcctggggg - cgttggcgtatagaaacatcctcgtaaagcgagaaggttcaagaatgtt	1091
Db	1011	cccaagcctggggg - cgttggcgtatagaaacatcctcgtaaagcgagaaggttcaagaatgtt	1070
QY	1092	cgttggagaaagaagccttcttccatgtgaatgagccggttgcagcttaagaagccggtgtcac	1151
Db	1071	cgttggagaaagaagccttcttccatgtgaatgagccggttgcagcttaagaagccggtgtcac	1130
QY	1152	ataaagcggtatcaaatggtgagcagctgtctgtacacagttgtgtacaaagcccaatagatggccct	1211
Db	1131	ataaagcggtatcaaatggtgagcagctgtctgtacaaagcccaatagatggccct	1190
QY	1212	gatgcacaaagtcttcgctcggaaagagcttaagcatcatcctcccaagcgagccaacgcgcga	1271
Db	1191	gatgcacaaagtcttcgctcggaaagagcttaagcatcatcctcccaagcgagccaacgcgcga	1250
QY	1272	caagccaggttgaatggcagatgcacaactcgacccctgcggtgtgaagaatgttactcggccca	1331
Db	1251	caagccaggttgaatggcagatgcacaactcgacccctgcggtgtgaagaatgttactcggccca	1309
QY	1332	gaagccacaagcggtgcacacccgc - cgaagaagcggaagaactctcaagttcctttagaacc	1390
Db	1310	gaagccacaagcggtgcacacccgc - cgaagaagcggaagaactctcaagttcctttagaacc	1369
QY	1391	cctaagctgttggagactgaatgaacggtgtggtcgagaagagagcggtcggatcttggggagggcc	1450
Db	1370	cctaagctgttggagactgaatgaacggtgtggtcgagaagagagcggtcggatcttggggagggcc	1429
QY	1451	tgaaagacttggcccccggtgcaaggagactct - ccagagctctcctccccccttgcagagccag	1509
Db	1430	tgaaagacttggcccccggtgcaaggagactct - ccagagctctcctccccccttgcagagccag	1489
QY	1510	caacatgtgtcccaagatgttggaaagggtcctcctcttcgcagatgttttgggtgtgtcat	1568
Db	1490	caacatgtgtcccccagatgttggaaagggtcctcctcttcgcagatgttttgggtgtgtcat	1549

QY	1570	ggggtctcccaaccactctctcaagtggttggtgagtcgagggagccaaccccaagctctctgc	1629
Db	1550	tggtgtccccaaccactctctcaagtggttggtgagtcgagggagccaaccccaagctctctgc	1609
QY	1630	cagatcaaccctcgagggcgacacctccaagccaatagttctcgggggtggtgagtcgagga	1689
Db	1610	cagatcaaccctcgagggcgacacctccaagccaatagttctcgggggtggtgagtcgagga	1668
QY	1690	ggcctatggttctcttgagagatctcctgcaacctcaagagactcccaagcgctcaagcgt	1749
Db	1669	ggcctatggttctcttgagagatctcctgcaacctcaagagactcccaagcgctcaagcgt	1728
QY	1750	ggatctgtctcctctgtgaggaacaagggtgcctaataatacatcttcgtcttaata	1808
Db	1729	ggatctgtctcctctgtgaggaacaagggtgcctaataataatacatcttcgtcttaata	1787
RESULT	6		
LOCUS	HSA276102	1428 bp	linear
DEFINITION	Homo sapiens mRNA for GPRC5C protein.		PRI 30-Jul-2000
ACCESSION	AJ276102		
VERSION	AJ276102.1	GI:9588668	
KEYWORDS	gprc5c gene.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Robbins,M.J., Michalovich,D., Hill,J., Calver,A.R., Medhurst,A.D., Gloger,I., Sims,M.A., Middlemiss,D.N. and Pangalos,M.N.		
TITLE	Molecular cloning and characterisation of two novel retinolic acid-inducible orphan G-protein-coupled receptors (GPRC5B and GPRC5C)		
JOURNAL	Genomics 76, 8-18 (2000)		
REFERENCE	2 (bases 1 to 1428)		
AUTHORS	Michalovich,D.		
JOURNAL	Direct Submission		
FEATURES	Submitted (22-MAR-2000) Michalovich D., Bioinformatics, SmtHkLine Beecham, Third Avenue, Harlow, Essex CM19 5AW, UK		
SOURCE	Location/Qualifiers		
gene	1. 1428		
CDS	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	103. 1428		
	/gene="gprc5c"		
	103. 1428		
	/gene="gprc5c"		
	/codon_start=1		
	/product="GPRC5C"		
	/protein_id="CAC00633.1"		
	/db_xref="GI:9588669"		
	/translation="MAIHKALVMGLPLPLFPANAGQGHVPPCSOGLNPIYNNLDD RSGWGVLEAVVAGAVITVFVLTITLIVASLPVQDTRKSLSTOVFLIGLTGLFC LVPFCVVKPDEFTSCASRFLFGLVFALCFSCGLAHVVALNLARKNGPRGWIPTVA LLLLELVETINTEWLITTLVSGSGEGPOGSSGMAVASFCAAVANDFWALLIYVIM LILAFGAPALPCGRKRMKHGVFLLTATGATVIAVMVYITVGNKOHNSPTMD DPTAIIAANAMAFVLFEVYIEVSOVTKSSPEGSOGDMYPRFGVGYENTLIKQKGO SMFENKAFSDMEVPAAKRPSPISGTINGOLITSVYPTKEALMHNKPPSECAVYIILP RATNSOVMSGSANSTFLRAEDYSAQSHQDAATPPDGKNSOVFRNPYVD"		
BASE COUNT	254 a 463 c 421 g 290 t		
ORIGIN			
Query Match	74.6%;	Score 1357.4;	DB 9;
Best Local Similarity	99.9%;	Pred. No. 1.1e-245;	Length 1428;
Matches 1358;	Conservative 0;	Mismatches 1;	Indels 0;
		Gaps 0;	
QY	49	gggaccacaacccagagccctgggcttgagagccagatagccatccacacaagaccttgtaag	108
Db	70	gggaccacaacccagagccctgggcttgagagccagatagccatccacacaagaccttgtaag	129
QY	109	tgcctgagactgcctctctctgtcttcccaaggagcctgggcccagggccatgtcccaacc	168

|||||
Db 130 TGCCTGGAGACTGCTCTCTCTGTTCCAGGGGCTGGGCCAAGGGCATGTCCACACC 189
Oy 169 ggtctgaagcaaggacctaaaccccttactaacaacctgtgtgacggtctcggggcgctg 228
Db 190 GGTGTACACCAAGGGCTTAACCCCTGTACTACAACTGTGTGACGGCTGTGGGGCGTGG 249
Oy 229 ggaatcgtctgagggcgctgtgtgagggagcatgttcacacacgttgtgtctacacatc 288
Db 250 GGCAATCGTCTGGAGAGGGCGGTGGGTGGGGCATGTGCACACGTTGTGTCTACACATC 309
Oy 289 atcctgtgtgcaagcctcccttgtgtgaggaacacaaagaaacggagcctgtgtggagac 348
Db 310 ATCTGTGTGGCGAGGCTCTCCCTTGTGTGACAGACAAAGAAACGAGAGCTGTGGGGAC 369
Oy 349 caggatctctctctgtgaggaacctgtgacctctctctcgtgtgtgtgtgtgtgtgt 408
Db 370 CAGGATATCTTCTCTGTGTGGGACCTGGGGCTCTTCTGCTGTGTGTGTGTGTGTGTG 429
Oy 409 aagccgaactctcgaactgtgacctctcgtcgtcgtcctctctgtgtgtgtgtgtgtgt 468
Db 430 AAGCCGCACTTCTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 489
Oy 469 tgtctctctgt 528
Db 490 TGTCTCTCTGT 549
Oy 529 gggcccgagggt 588
Db 550 GGGCCCCGGGGTGGGTATCTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 609
Oy 589 atcaataaagatgt 648
Db 610 ATCAATACAGATGT 669
Oy 649 ggaacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 708
Db 670 GGCAACACACGAGCGGTGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 729
Oy 709 gtcatgagcaccatctagatgcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 768
Db 730 GTCATGGACATCATCTAGCTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 789
Oy 769 gccctgt 828
Db 790 GCCCTGT 849
Oy 829 gcaacctcgttgtgcatagt 888
Db 850 GCCACCTCCGTTGGCATATGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 909
Oy 889 cacaacagttccacactgt 948
Db 910 CACAACTGT 969
Oy 949 gccctgt 1008
Db 970 GCTTGT 1029
Oy 1009 caaagctcaggggt 1068
Db 1030 CAAAGCTACAGGGGGGATGTATACCCACCGGGGGCGGTGTGTGTGTGTGTGTGTGT 1089
Oy 1069 gaggcaagaggtcagagcagttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1128
Db 1090 GAGCAAGAGGGTCAAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1149
Oy 1129 gcaagtaagagcggt 1188
Db 1150 GCAGCTAAGAGCGGGT 1209
Oy 1189 taccagcccaactgt 1248

Db 1210 TACCAGCCCACTGAGATGGAGCCCTGTATGTGCACAAAGTTCCTCCGGAAGGACTTACGACATC 1269
Oy 1249 atcctcccaagggccacacgcagcaacagcagggtgtgtgtgtgtgtgtgtgtgtgtgtgt 1308
Db 1270 ATCTCTCCACGGCCGACCCGACCAACGACGAGGTGATGGCAATGTCCAACTCCACCTGTGG 1329
Oy 1309 gctgaagacatgtactcgtgcccagagccacagcgctggtgacacacgcgcgaaagagcgcaag 1368
Db 1330 GCTGAAGACATGTACTGTGGCCAGAGCCAGACGAGGGGCTACACCGCGGAAGAGCGCAAG 1389
Oy 1369 aactcaggtccttagaaccctcagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1407
Db 1390 AACTCTCAGCTTTTAGAAACCCCTACGCTGTGGAGCTGA 1428

RESULT 7
AX014742
LOCUS AX014742 1326 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from Patent WO953054.
ACCESSION AX014742
VERSION AX014742.1 GI:10041009
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 1326)
AUTHORS Michalovich D., Hill J., Medhurst A. and Pangalos M.
TITLE Axord, a protein-coupled receptor
JOURNAL Patent: WO 953054-A 1 21-Oct-1999;
SMITHKLINE BEECHAM PLC (GB)
FEATURES
source 1..1326
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 232 a 431 c 386 g 277 t
ORIGIN

Query Match 72.8%; Score 1324.4; DB 6; Length 1326;
Best Local Similarity 99.9%; Pred. No. 1.8e-239;
Matches 1325; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 82 atggccalcacaaagcctgt 141
Db 1 ATGGGCATGCCAAGAGCTGT 60
Oy 142 gccctgt 201
Db 61 GCTTGT 120
Oy 202 aacctgt 261
Db 121 AACCTGT 180
Oy 262 atgttcacaaagctgt 321
Db 181 ATTGTACACAGCTTGT 240
Oy 322 accaagaaagagcgt 381
Db 241 ACCAAGAAACGAGCCTGT 300
Oy 382 ttcgtcctcgt 441
Db 301 TTTCCTCTCGT 360
Oy 442 ttcctctgt 501
Db 361 TTTCCTCTTGT 420
Oy 502 ctcaactctgt 561

Db	421	CTCAACTTCCTCGGCGCCGGAAGAACCAGCGGGCCCGGGGCTTGGTGATCTTTCACATCTGGCT	480
QY	562	ctgctgctgacccctggtgtagagtgatcatcaatacagaagtgtgcatcatcaccttggt	621
Db	481	CTGCGTGTGACCCCTGTGTAGAGTATCATCATACAGAGTGGTGTATCATCACCTGGT	540
QY	622	cggggcagtgagggagggcgccctcaaggcacaagcagcagcttgggccgttgctcc	681
Db	541	CGGGGCAATGGCGAGGGGGGCCCTCAGGGCAACACAGCGCAGGCTGGGGCGTGGCTTC	600
QY	682	cccgtgtgcafcgcaacaatggaacttgtaatgacacatactagatgctgtgtgt	741
Db	601	CCCTGTGCGCCGTCGCCACATGGACTTTTTCATGTGCACATCATCTACGTATGCTGTG	660
QY	742	ctgggtgtgctcttcgggggagcttggcccgccctgtgtgtgacgagctacaagcgtctgctaa	801
Db	661	CTGGGTGCTTTCCTGTGGGGCGCTGGGCCCGCTGTGTGGCCGCTACAAACGGTGGGTAA	720
QY	802	catgggtgtcttctgtgtctctcaacagccacccctcgttgcataatgggtgtgtgac	861
Db	721	CATGGGGCTTGTGTGCTCTCTCACCAACCCACCTCCGTTCATATGGGTGTGTGATTC	780
QY	862	gtcaatgtaacttaagggaacaagaagacaacaagctccacacttgatgtagcccaagctg	921
Db	781	GTCATGTACTTACGGACACAGACAGACACATCTCCACCTGGGAATGACCCAGCTG	840
QY	922	gcaatcgcccttcgcgcgaatgatcctggggccttcgtctctcttctagatcatcccgagctc	981
Db	841	GCCATGCGGCTTCGCGGCCAATGCTGGGGCTTTCGTCTCTTCTACGTCTACCCGAGGTC	900
QY	982	tcccaagtgaaccaagctcagcccagaagaagctacagagggagacatgtatcccaacgg	1041
Db	901	TCCCGAGTGACCAAGTCCAGGCCACAGCAAAAGCTACAGGGGGACATGTATACCCACCGG	960
QY	1042	ggcggtgggctcttgagaccatccttgaagaagaagaaggtctagaagatgtctgtgaagaac	1101
Db	961	GCGGTGGGCTTATGAGACCATCTCTGAAGAAGCAAGAAGGTCAGAGCATGTCTGTGAGAAC	1020
QY	1102	aagagccttccatgcatgtagagcggtctgcagctaaagagcggtgtcaacatatacgcg	1161
Db	1021	AAGGCTTTTTCATGTGATGAGAGCGGTTGTCAGTTAAGAGGCCGCTTCACATATACGCGG	1080
QY	1162	tacaaatgggcagctgtctgaccaaigtgttaccagcccaatgagatggccctgattgcacaaa	1221
Db	1081	TACATATGGGAGCTGCTGACCAAGTGTGTATACAGCCCATGAGATGGCCCTATGCACAAA	1140
QY	1222	gttcgcgtccgaaggagctttagacatcatctcccaaggggccaacggccaagaagcaggtg	1281
Db	1141	GTTCGCTCCGAAGGAGCTTACACATCATCTCCACGGGGCCACCGCCAAACGCGAGGTG	1200
QY	1282	atgggcagtgcaaatctgacccctgagggtctgaagacatgtactctggcccaagaccag	1341
Db	1201	ATGGGCATGTGCCAATCTGACACCTCGGGGCTGAAGACATGTACTCGGCCACAGGCCACCA	1260
QY	1342	gcggccacacgcgcgaaagaagcgaagaactctcagctcttaagaacccctacgtgtgg	1401
Db	1261	GGGGCCACACCGCCGGAAGAGCGGCAAGAACTCTGAGGCTTTTGAAGAACCCCTACGTGTGG	1320
QY	1402	gactga 1407	
Db	1321	GACTGA 1326	
RESULT 8			
AX083413			
LOCUS	AX083413	1323 bp	DNA
DEFINITION	Sequence 105 from Patent WO0112660.	linear	PAT 28-FEB-2001
ACCESSION	AX083413		
VERSION	AX083413.1	GI:13185250	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE	1 (bases 1 to 1323)	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	Kato, S. and Kimura, T.	
TITLE	Human proteins having hydrophobic domains and dna encoding these proteins	
JOURNAL	Patent: WO 0112660-A 105 22-FEB-2001;	
FEATURES	SAGAMI CHEMICAL RESEARCH CENTER (JP) ; Proteogene Inc. (JP)	
source	Location/Ouallifiers 1. .1323 /organism="Homo sapiens" /db_xref="taxon:9606"	
BASE COUNT	232 a 431 c 384 g 276 t	
ORIGIN		
Query Match	72.7%; Score 1323; DB 6; Length 1323;	
Best Local Similarity	100.0%; Pred. No. 3.2e-239;	
Matches 1323;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	82 atgagccatcacaaagacctggtgtagtgcctgaggactgcctctctctgctccagg 141	
Db	1 ATGGCCATCCCAAGACCTTGATGTGGCCGGAGACGCCCTCTCTGTCCTCCAGG 60	
QY	142 gctctggcccgaggccatgctccaccggctgcagaccgaaggtctaaacccctgtactac 201	
Db	61 GCTTGGGCCCAAGGCGCATATGCCACCGCGTGCACCAAGGCTCAACCCCTGTACTAC 120	
QY	202 aacctgtgtaacgctctcgaggcgctgaggcatcgtccctggaaagcgtgctggggcgagc 261	
Db	121 AACCTGTGTGACCGCTGTGGGGCGTGGGGCATATGCTTGGAGGCCGTGTGGGGCGGC 180	
QY	262 atgtcacccagcttgtgtgtcaccatcatctctgtgtgcagcgtccctcttltgcaagac 321	
Db	181 ATTTGCACACAGTTGTGTGTACATCATCTGTGGCCAGCGTCCCTTTGTGTGAGAC 240	
QY	322 accaagaaacggagcctgtgaggaaaccaagtatctctctctgtggagacctgtggctc 381	
Db	241 ACCAAGAAACGAGCGCTGTGGGGACCCAGGATTTCTTCTTGGGACCCCTGGGCCCTC 300	
QY	382 tcttgccctcgtgtgttcctgt 441	
Db	301 TTCTGCTGT 360	
QY	442 ttcctcttctgt 501	
Db	361 TTTCCTTTTGGGGTCTGT 420	
QY	502 ctcaactctctgtgcccggagaagaaacagggcccgagggtgtgtgtgtgtgtgtgtgtgtgt 561	
Db	421 CTCACATTCCTGTGGCCGGGAAGAACACGCGGCCCGGGGGTGGGTATCTTCACTGTGGCT 480	
QY	562 ctgctgt 621	
Db	481 CTGCTGT 540	
QY	622 cggggcagctgtgagaggcgccctcaaggcgaaacagcagcgagctgtgagcgctgtgctc 681	
Db	541 CGGGCGAGTGGGGAAGGGCGGCCCTCAAGGCAACAGCGAGCGTGGGGCGTGGCGCTTC 600	
QY	682 cccctgtgcacatgcacaacatgtgacttgtcatgtgacatcatctagcgtgactgtgtgtgt 741	
Db	601 CCTGTGCGCATGCGCAACATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 660	
QY	742 ctgagtgctctctctgt 801	
Db	661 CTGGGTGTCT 720	
QY	802 catgggtgtcttgt 861	
Db	721 CATGGGTGTCTTGT 780	
QY	862 gtcatgtacttactacgcaaacagcagcatcccaactgtgatatcccaagctgtgtgtgtgtgtgtgt 921	

```

Db 781 CTCATGTATACCTTACGGCACAACACAGACAGTCCACCTGGATGAGTACCCACGCTG 840
Oy 922 gcaatgocctcgccgcgaatgctgagcctgctctctctcaatgcatcccaagctc 981
Db 841 GCCATGCCCTCGCCGCCCAATGCTGGCCCTTCTCTCTTACGTATCTCCCGAGGTC 900
Oy 982 tcccaagtgacaagctcgaagcccaagcaaaagctcacaagggagacatgtaaccacccg 1041
Db 901 TCCAGAGTACCAAGTCCAGGCCAGAGCAAAAGTACAGGGGGAGCATGTACCCACCCGG 960
Oy 1042 ggcgtgagctatgagacatctctaaagagagagagagagagagagagagagagac 1101
Db 961 GGCCTGGGCTATGAGACCATCTGAAGAGAGAGAGGTCAGACATGTTGTGAGAGAC 1020
Oy 1102 aagcccttcacatgagatgagccggttgacagtaagagagccggtgtcaccaacagcgg 1161
Db 1021 AAGGCTTTTTCATGATGATGAGACCGGTTGACACTAAGAGCCGGGTGTACCATACAGCCGG 1080
Oy 1162 tacaatggcagctgctgacacagtgctgtaaccagccacatgagatgacccatgacacaa 1221
Db 1081 TACAATGGCAGCTGTGACAGAGTGTGACAGCCCACTGAGATGGCCCTGATGCACAAA 1140
Oy 1222 gttccgtccgaagagctttagacataatctctcccaagggccacccgaagcaggtg 1281
Db 1141 GTTCCGTCCGAAGAGCTTACGACATCATCTCCACGGGCCACCGCCACAGCAGGTG 1200
Oy 1282 atggcagctgcaactgacacccctgagggctgaagacatgtaactgagcccaagcaccag 1341
Db 1201 ATGGGCACTGCACTGACCTCGGGCTGAAGACATGTACTGCGCCAGAGCCACAGCAG 1260
Oy 1342 gcggcacaacccgcgaagacgagcaagactcgaagctttagaaccctcagctgtg 1401
Db 1261 GCGGCAACACCCGCCGAAGAGAGAGCAAGACTGAGCTTTAGAAAACCCCTACGCTGTG 1320
Oy 1402 gac 1404
Db 1321 GAC 1323

```

```

RESULT 9
AC079325/c 192087 bp DNA linear HTG 09-FEB-2002
LOCUS Homo sapiens chromosome 17 clone RP11-19403 map 17, ** SEQUENCING
DEFINITION AC079325
IN PROGRESS ** 5 unordered pieces.
ACCESSION AC079325.5 GI:18640666
KEYWORDS HNG: HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 192087)
Birten,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-19403
Unpublished
2 (bases 1 to 192087)
Birten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bede,F., Boguslavsky,L.,
Boukhalter,B., Brown,A., Burkett,G., Campolongo,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heatford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Larocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G.,
Macdonald,P., Margulis,N., McCarthy,M., McKean,P., McKernan,K.,
McPhaeters,R., Meldrum,J., Menus,L., Mihova,T., Mlenka,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,

```

Tirelli,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-AUG-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 8, 2002 this sequence version replaced gi:117975276.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence-submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10900
Center clone name: 194_J_3

* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 13221: contig of 13221 bp in length
* 13222 13321: gap of 100 bp
* 13322 116001: contig of 102680 bp in length
* 116002 116101: gap of 100 bp
* 116102 130025: contig of 13924 bp in length
* 130026 130125: gap of 100 bp
* 130126 162945: contig of 32820 bp in length
* 162946 163045: gap of 100 bp
* 163046 192087: contig of 29042 bp in length.

FEATURES
Location/Qualifiers
1..192087
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-19403"
/clone_lib="RPC1-11 Human Male BAC"

BASE COUNT 49249 a 47779 c 47513 g 46919 t 627 others
ORIGIN

Query Match 59.8%; Score 1088; DB 2; Length 192087;
Best Local Similarity 99.5%; Pred. No. 3.3e-195;
Matches 1091; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

Oy 39 tcaacctggaagaccccaagacagcctgctggaagcagagatgagccacaaagc 98
Db 156609 TCTGCTCTTAGGAGACCAACAGAGCCTGGCTGGAGCCAGAGTGCATCCAAAGC 156550
Oy 99 ctgtgatgtgcttgagactcctctctctctctctctctcctcctcctcctcctcctc 158
Db 156549 CTGTGATGATGTCCTGGAGACATGCTCTCTCTGTTCCAGAGGCGCTGGGCCA 156490
Oy 159 tctccacccgctgacagcaagagcctcaacccctctactacaactgtgtacgcctc 218
Db 156489 TGTCCACCCCGCTGACAGCCAAAGCCCTCAACCCCTGACTACAACTGTGTGACCCCTC 156430
Oy 219 tggagcgttggagcctgtctccctgagagccgtgtgctggggagcggcattgtcaccaagttgt 278
Db 156429 TGGGGCGTGGGCAATCGCTCGAGGCGCTGGCGGGCGGCAATGTGCACACCTTTGT 156370
Oy 279 gctacacatcctctgtgtgacagcctccctctgtgtgaagacacaaagagagcct 338
Db 156369 GCTACACATCATCTGTGTGGGACAGCTCTCTTGTGTGAGAGACACAAAGAGAGAGCT 156310
Oy 339 gctggagaccaggtatctctctctctctgtggagccctgggacctctctgtgctgttgc 398

```


[illegible]

Db	9418	CTTGGAGCCAGATGGCCACTCACAGAACTTTGCTGATGTGCTTAGACTGCTCTCTT	9359
Qy	129	ccctgtcccaaggagccctgcccacagagccatgtcccaaccgagctgtcaagccaaggccctcaa	188
Db	9358	C---TTCCCAAGAGCCTTGGCCCAAGAAATATGCCCCACCTGGTTGCACCCCAAGATCTGGA	9302
Qy	189	cccccttactaanaacctgtgtaacgcctctcgtgggcgttgaggcatcgtctctgaagccgt	248
Db	9301	TCCCTCTACTACAACTCTGTGACCGCTCGGGGGGCGTGGGCGATTTGCTTGGAGGCGAGT	9242
Qy	249	ggctgggggggggcatlgttaaccaagtttgtcttaaccaatcaatccctggtggccaacccccc	308
Db	9241	GGCTGGAGCGCATCATCGACATTTTGTATTAACATATCATCTTGTGGGTACCTTC	9182
Qy	309	cttgtgtcaagacaaccaagaacgagagcctgtcgtggagcaaccagatctctctctctcgtgg	368
Db	9181	ATTGTGTGAGACACTTAAGAGGGAGGAGCCTCTTGGGAGACCAAGTGTCTTCTGTCTGGG	9122
Qy	369	gaaccctggagccctctctcgtccctgtgttgcgcgtgtgtgtgaagcccgaaactctccacgt	428
Db	9121	CACCCCTGGGTCTCTTTCGCTCGTGTGTTGCTGATGTTGGTGAAGCGGAGCTTCTACCTG	9062
Qy	429	tgcctctcggagcctctctctcttcttggggtctgttgcgcaatctctctctctgtctcggcgc	488
Db	9061	TGCCTCTGAGAGCTTCTCTTTTGGGGTCCGTGTTTGGCATCTCTCTCTCTCTGTCTGTGAGC	9002
Qy	489	tcaagctcttggccctcaaaactctctcgtcccgagaagaaccaacgagcccgaggcgtgtgtat	548
Db	9001	TCAGCTCTTTCCTTCAACTTCTTAACCCGGAAGAACATGAGGCCCAAGGCTGGGTGAT	8942
Qy	549	cttcaactgtgctctgtctgtctgaacctgtgtagaagtcatcatcaatlaagaaagtgtcgtat	608
Db	8941	CTTTCACGATGGCGCTGCTGCTCAACCTTGGAGGATCATTAACACCGAGTGAGCTTAT	8882
Qy	609	cataacctgtgtcgtgggaggtgtgcgaaggcggccctcaaggcaacagcagcgcagcgtcgt	668
Db	8881	CATCACCTGTGTAGGGGAGGTGGCCAGGTTAGCCCCCGGGCAATGTCACTGCGAGATC	8822
Qy	669	ggccgttggacccccccctgtgcacgtgcacaatagagacttgcacgtgcaaccatacctagt	728
Db	8821	GACCATGACCTTCTCGTGTGCATCGGCANACATGAGACTTGTGATGAGGCTCTCATCTAGCT	8762
Qy	729	catgctgctgctgtctgtgtgagtgacctctctcgtggggccttgcgcgcgtctaa	788
Db	8761	AATGCTGTGTGCTGTGACGGGCTTCCATAGAGGCTGGCCACCTTGTGTGGGCGGCTTCAA	8702
Qy	789	gcgctgacgctaaagcatgggggtcttgtgtctctcacacagcgaacctcogtltgcatatg	848
Db	8701	GCGGTGGGGGAAACACGGGGCTTTTGTGTGTGCACACCATGTCTCATCTGCGCATCTG	8642
Qy	849	ggtgtgtgtatgtgtatgtataacttaaggcaaacagagcagcaaacagttcccaactgtga	908
Db	8641	GGTGTATGAGTTTCTCATGTACACTTAACGGCAACGACACACATAGCCCACTTGGA	8582
Qy	909	tgaaccacgctgtgcacatcgccctcgcgcgcgaatgctgtggagctctcgtctctcttaagt	968
Db	8581	TGACCCCAACGTTGGCCATTTGCCCTCGCTGGCATGCTGTGAGACTTTGTCTCTTAGT	8522
Qy	969	catccccaaggtctcccaagtgacacaagtlccagccacagagaagaactacaaagggagat	1028
Db	8521	CATCCTGTAGGTTCTACAGAGGTGACCAAACTTACCTCCAGAACAGACCTTACCAAGGGGACAT	8462
Qy	1029	gtaacccaacccggggcgltgggagctatagaccatccttgaaagaagcagaaggttcaagaagt	1088
Db	8461	GTACCCGACCCGAGGGGTGTATGTAGACCATCTGAAAGGAGCACAGCGGGCCAGAGCAT	8402
Qy	1089	gtctgtgagaacaaggcctcttccatgtatgtgaagccgagtttgagct	1134
Db	8401	GTTTGTGAGAACAAAGCATTTTCTATGTAGACGAACACAGCCTCAAGGT	8356
RESULT 12			
AL669969/c			

RESULT 12
AL669969/c

LOCUS	AL669969	196975 bp	DNA	linear	HTG 30-JAN-2002
DEFINITION	Mus musculus chromosome 11 clone RP23-254J18, *** SEQUENCING IN PROGRESS ***. In unordered pieces.				
ACCESSION	AL669969				
VERSION	AL669969.3	GI:18477062			
KEYWORDS	HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.				
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	1 (sites)				
TITLE	Plumb.B.				
JOURNAL	Direct Submission				
	Submitted (29-JAN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Feb 1, 2002 this sequence version replaced gi:18375930.				
COMMENT	----- Genome Center Center: Wellcome Trust Sanger Institute Center code: SC Web site: http://www.sanger.ac.uk Contact: humquerry@sanger.ac.uk ----- Project Information Center project name: BM254J18 ----- Summary Statistics Assembly program: XGAP4; version 4.5 Sequencing vector: plasmid; 108752; 100% of reads Chemistry: Dye-terminator Big Dye; 77% of reads Chemistry: Dye-terminator; 22% of reads Consensus quality: 193875 bases at least Q40 Consensus quality: 194270 bases at least Q30 Consensus quality: 194637 bases at least Q20 Insert size: 195775; sum-of-contigs Insert size: 185198; agarose-fp Quality coverage: 11.57x in Q20 bases; sum-of-contigs Quality coverage: 12.31x in Q20 bases; agarose-fp ----- * NOTE: This is a 'working draft' sequence. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. ----- Location/Qualifiers 1. 196975 /organism="Mus musculus" /db_xref="taxon:10090" /chromosome="11" /clone="RP23-254J18" /clone_lib="RPCT-23" 1. 5154 /note="assembly_fragment:02093 fragment_chain:1 clone_end:SP6 vector_side:left" 5255. 35776 /note="assembly_fragment:01566 fragment_chain:1" 35877. 55304 /note="assembly_fragment:01716 fragment_chain:1" 53405. 66379 /note="assembly_fragment:00116 fragment_chain:2" 66480. 82119 /note="assembly_fragment:02883 fragment_chain:2" 82220. 88071 /note="assembly_fragment:03245 fragment_chain:2" 88172. 108526 /note="assembly_fragment:00436 fragment_chain:2" 108627. 115927 /note="assembly_fragment:04402				

	misc_feature		fragment_chain:2"
		116028:	.133671
		/note="assembly_	fragment:02982
	misc_feature		fragment_chain:2"
		133772:	.144593
		/note="assembly_	fragment:02764
	misc_feature		fragment_chain:2"
		144694:	.167710
		/note="assembly_	fragment:02179
	misc_feature		fragment_chain:2"
		167811:	.183647
		/note="assembly_	fragment:03138
	misc_feature		fragment_chain:2"
		183748:	.196575
		/note="assembly_	fragment:03589
		fragment_chain:2	
		clone_end:77	
		vector_side:right"	
BASE COUNT	51273	a	47120 c 46010 g 51360 t 1212 others
ORIGIN			
Query Match	45.5%	Score 826.8;	DB 2; Length 196975;
Best local Similarity	84.0%;	Pred. No. 4.7e-146;	
Matches 946;	Conservative	0;	Mismatches 177; Indels 3; Gaps
QY	9	gccctcaaccacccggaagtagcagtcgcgtccagcccgagggagaccacaagaccttg	68
Db	131703	GACCTTGAGAGACGATTCTGTCTGTCTTCCCTCTCTATAGGACCACGACTGAACCTGG	131644
QY	69	cctggagacaggatlggcacacccaagaaccttggatgatgcatcctggagactcctct	128
Db	131643	CCTGGAGACCGAGATGGCCCACTCACAAACCTTGATGTGCCTAGAGACTGCCCTCTT	131584
QY	129	ccgtttccccaggagccctgggcccaggcccatgtcccaaccggcctgcagcccaaagcccaa	188
Db	131583	C---TTCCCAGAGCCTTTGGCCCAAGAATTCATGCCCCCACTGGTTGACGCCACAGTTCGA	131527
QY	189	ccccctgtacatacaacctgttgtacgcqctctggggcgctggagcactgcctcggagacgt	248
Db	131526	TCCCTCTACTACAACCTCTGTGACCCCTCGGGGGCCTGGGGCATTTCTCTTGGAGGCAAT	131467
QY	249	gacctgggagcgagcatgtcaccaacgtttgtgtcaccaatccctgglyggcaagcctccc	308
Db	131466	GGCTGGAGACGACATCATCAAGACATTTGGTATAACCATCATCTCTTGCGCTAACCTTCC	131407
QY	309	ctttgttcagagacaccaagaagaaagacccgtctggggagccagtatcttcctcctctggg	368
Db	131406	ATTGTGTGAGGACACTAAGAAAGCGGAGCCTCTTGGGGACCCAGGATGTTCTTCCTGCTGGG	131347
QY	369	gacctggagccctctctgtccctctgcgtgtgttcgctgtgtgttaagccgaacttccaacctg	428
Db	131346	CACCTTGGGTCTTCTTGCCCTCGCTTTGGCTGTGTGTGAAGCGGACTTCTCTACTCTG	131287
QY	429	tgcctctcggcgtcttccctccttcttggggattctgtgtccaatcgtcttcttctgtctgcgcg	488
Db	131286	TGCCCTTCGAGAGCTTCCCTTTGGGGTCTGTTGGCATTCGATCGCTTTCCTGCTGGTAGC	131227
QY	489	tcaagctcttgcctcaactctctgcgcgggaagaaccaagggcccggggctcggatgat	548
Db	131226	TCAAGCTCTTTCCTTCACATCTTCCATMACCGGAAAMACATAGGGGCCCGAGGCTGGGTAT	131167
QY	549	cttacatgttgtctgtctgtctgagaccggttagaagtgcatactaatacaagatggtcgtat	608
Db	131166	CTTACCGTGGCGCTGCTGCATCACCCCTTGTGGAGGTATATATTAAACCCAGAGGCTTAT	131107
QY	609	catcaacctgttctggggcagctggcgaaggcgcccccacaaggcaacagcagcgacagctg	668
Db	131106	CATACCCCTGGTACGGGGAGGTGGCCAGGTTAAGCCCCCTGGGCATATAGTCCGCACTC	131047
QY	669	ggccgtggacctccccctgtgcacatgcgcaaatygacctgtgtcalatgagcatactacgt	728
Db	131046	GACCATTAACCTCTCGTGTGCATCGCCAAATGAGATTGTGTATGGCTCTCATTTACGT	130987

LOCUS	AX136494	613 bp	DNA	linear	PAT 30-MAY-2001
DEFINITION	Sequence 416 from Patent EP1067182.				
ACCESSION	AX136494				
VERSION	AX136494.1	GI:14272898			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. (bases 1 to 613)				
AUTHORS	Ota,T., Isoigal,T., Nishikawa,T., Kawai,Y., Sugiyama,T. and Hayashi,K.				
TITLE	Secretory protein or membrane protein				
JOURNAL	Patent: EP 1067182-A 416 10-JUN-2001;				
FEATURES	Helix Research Institute (JP)				
source	location/Qualifiers				
	1..613				
BASE COUNT	90 a 200 c 175 g 143 t				5 others
ORIGIN					
Query Match	31.7%; Score 575.8; DB 6; Length 613;				
Best Local Similarity	99.1%; Pred. No.1.5e-98;				
Matches 577; Conservative	0; Mismatches 5; Indels 0; Gaps 0;				
QY	16 ccagccggaagatcagatcgcgtcagccttggagggagagcccaacagagcctggcttggga	75			
Db	20 CAAGCGGAAATGATGAGAGTGGGCTCAGCTCGAGGAGGAGCCACAGAGCTGGCTGGGA	79			
QY	76 gccagatgagccatccaaaagccttgatgatgtgcctggagcttcctctctctctctc	135			
Db	80 GCCAGAGATGGCCATCCACAAAGCCTTGATGATGCTCGGAGCTCCTCTCTCTCTCTCTCT	139			
QY	136 ccaaggagccttggcccaaggcattgccaccggccttgaagcgaagagctcaaccctctg	195			
Db	140 CCAGGGGCGCTGGGCCAGAGGCGATGCCCAGCCGGCTGAGCCAAAGGCGCTCAACCCCG	199			
QY	196 tactcaaaccttgtagccgcctctggggcgtagggacatgctcttggagagcgtagctggg	255			
Db	200 TACTCAACCTTGATGAGCGCTCTGGGGCGTGGGCGATGCTCTGAGGCGCTGGCTGGG	259			
QY	256 gggggcatgtcacacagcttctgtcacacacatcctgttggcagagctccctctctgtg	315			
Db	260 GCGGGGATTTGTACACACAGCTTTGTGCTACACATCANTCTGTTGGCCAGCGCTCCCTTTGG	319			
QY	316 cagagacacaaagaacagagcctgcttggggaccagagtaattcttcctcttggagacctg	375			
Db	320 CAGGACACCAAAAGAGGAGCTGCTGGGGAGCCAGGTAATTTCTTCTTGAGGAGCCCTG	379			
QY	376 ggcctctcttgcctcgtgtcttgcctcgttggatggtagagccagaccttctcaacctgtacct	435			
Db	380 GGCCTCTTCTGCTCTGTTTGCCTGTGTGGTGAAGCCCGACTTCTCCACCTGTGCTCT	439			
QY	436 cggagcctctctcttggagttctgtctgcacatctgctcttccttcttcttggcagctc	495			
Db	440 CCGGCGCTTCTCTTGGGGATTTGTATGCGCATCTGCTTCTTGTGCTGGCGGCTCAGGTC	499			
QY	496 ttgtgcctcaaatctcttggccgggagaaacacacggggcccggggcttggatgtaattcact	555			
Db	500 TTTTGGCCTTCACTTCTCTGGCCGGGAAGAACAGGGGCGCCGGGGCGGTGATCTTTCAC	559			
QY	556 gtggcctcgtctgtgaccttggtagagttcatcatcaatataca 597				
Db	560 GTGGATCTTCTCTGTGACCTTGATATAGGTATCATCATCANTACA 601				

LOCUS	AK000249	1845 bp	mRNA	linear	PRI 22-FEB-2000
DEFINITION	Homo sapiens cDNA FLJ20242 fis, clone COLF6369.				
ACCESSION	AK000249				
VERSION	AK000249.1 GI:7020202				
KEYWORDS	Oligo capping; fis (full insert sequence).				
SOURCE	Homo sapiens colon mucosa cDNA to mRNA, clone_11b:COLF clone:COLF6369.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Matanabe,K., Kunagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.				
TITLE	NEDO human cDNA sequencing project				
JOURNAL	Unpublished (2000)				
AUTHORS	2 (bases 1 to 1845)				
TITLE	Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-FEB-2000) Smino Sugano, Institute of Medical Science, University of Tokyo, Depntment of Virology; Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)				
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- 6' 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).				
FEATURES	Location/Qualifiers				
source	1..1845				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="COLF6369"				
	/tissue_type="colon mucosa"				
	/clone_1b="ColF"				
	/note="Cloning vector PME18SFL3"				
	60..548				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="BA91030.1"				
	/db_xref="GI:7020203"				
	/translation="MAIHKALVWCLPLFLPFQSYOCMHYRFGVETFLKEKGQSMVEVKAFSMDPEVAKRVPSPYSYNQDLTYSYQPEMLMKVBEQAYDILLPRAVANSQVMSANSTLRADDMYSASOSHQAETPKDKNSVSELTORQDANTNOVFLRL"				
CDS					
BASE COUNT	391 a 569 c 519 g 366 t				
ORIGIN					
Query Match	20.8%; Score 378; DB 9; Length 1845;				
Best Local Similarity	100.0%; Pred. No. 2,2e-61;				
Matches 378;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1002 cccaagaagagctcccgagggaacatgttaccacccccggggcgctgagctatgagaccat 1061				
DB	113 CCCAAGCAAAACTVACCAAGGGGGAATATTACCCACCCGGGGCGTGGCTATGTGAGACCAT 172				
QY	1062 ccttgaagaagcagaaggttcagagcattctctgtgagaagacaagaagcctttccatgta 1121				
DB	173 CCTGAAGAAGCAGAAGGCTCAGAGCATTTTGCTGAGAGAACAAAGCCTTTTTCATGANTGA 232				
QY	1122 gccggtctgcagctaaagagccggtgtcaccatacagcgggtgtacaaatggcagctgtgac 1181				
DB	233 GCCGGTTCAGCTAAGAGCGCGGTGTCCACATACAGCGCGGTACCAATGGCGAGCTGTGAC 292				
QY	1182 cagttgtatcaccaagcccatctgagatgagcccttgatgcacaagaatttcgctcgaaagagctta 1241				
DB	293 CAGTGTGTACCAAGCCCATCTGAAGATGGCCCTTATGTCACAAAGTTCCGTCCGAAGAGCTTA 352				

```
QY 1242 cgacatcatcctcccaaggccacgccaagccagtgatgagtgagtgccaactcgac 1301
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 353 CGACATCATCTCTCCACGGGCCACCGCCAAACAGGTGATGGGCACTGCGAC 412
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1302 cctgcgggctgaagacatgtactcgcccaagccacccagcgccacccgcaaga 1361
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 413 CCGCGGGGTGAAGACATGTAATCGGCCAGAGCCACAGGGCGCCACCGCCGAAGA 472
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1362 cggcaagaactctcaagt 1379
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 473 CGGCAAGAACTCTCAGGT 490
    ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Search completed: September 22, 2002, 17:16:01
Job time: 9454 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 16, 2002, 09:26:51 ; Search time 31.69 seconds
(without alignments)
1545.710 Million cell updates/sec

Title: US-09-895-686-1

Perfect score: 2326

Sequence: 1 MAIHKALVMCLGLPLFLPPG.....ATPPKDGKNSQVFRNPYVWD 441

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_032802.*

- 1: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS1/gcgdata/hold-genseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2326	100.0	441	21	AAV57283 Human GPCR protein
2	2326	100.0	441	21	AAV50932 Human fetal brain
3	2326	100.0	441	22	AAV39849 Human polypeptide
4	2326	100.0	441	22	AAB88591 Human hydrophobic
5	2326	100.0	441	22	AAB88359 Human membrane or
6	2326	100.0	486	22	AAU14166 Human novel protei
7	2325	100.0	441	20	AAV49156 G-protein coupled
8	2320	99.7	460	22	AAW41635 Human polypeptide
9	2226	95.7	479	21	AAV44273 Human Metabotropic
10	2027	87.1	400	20	AAV36302 Human secreted pro
11	2004.5	86.2	461	20	AAV49157 G-protein coupled

12	1221	52.5	231	22	AAU14402	Human novel protei
13	1200.5	51.6	296	21	AAV44274	Mouse Metabotropic
14	1020.5	43.9	302	21	AAV50933	Human fetal brain
15	733	31.5	403	20	AAV32141	Human G-protein co
16	733	31.5	403	21	AAB43085	Human ORFX ORF2849
17	733	31.5	403	21	AAV57287	Human GPCR protein
18	733	31.5	403	22	AAV78615	Human G-protein SEQ
19	733	31.5	427	20	AAV32142	Human G-protein co
20	732	31.5	403	22	AAV93548	Human polypeptide,
21	558	24.0	357	22	AAV93311	Human protein sequ
22	558	24.0	357	22	AAB68891	Human RECAP poly
23	550.5	23.7	313	22	AAB29910	Peptide #2561 enco
24	550.5	23.7	313	22	AAB35088	Peptide #2594 enco
25	550.5	23.7	313	22	AAB20506	Protein #2505 enco
26	550.5	23.7	313	22	AAV55909	Human brain expres
27	550.5	23.7	313	22	AAV68279	Human bone marrow
28	550.5	23.7	313	22	AAV16102	Peptide #2536 enco
29	550.5	23.7	313	22	AAV28593	Peptide #2630 enco
30	550.5	23.7	313	22	AAV03829	Peptide #2511 enco
31	533	22.9	347	21	AAV76071	Rat skin cell tran
32	533	22.9	347	22	AAB56010	Skin cell protein,
33	524	22.5	317	22	AAB12274	Human orphan GPCR
34	524	22.5	362	22	AAB06764	Human G-protein co
35	518	22.3	323	22	AAB69174	Human G-protein co
36	478	20.6	106	20	AAV36337	Human secreted pro
37	347	14.9	200	20	AAV76571	Human ovarian tumo
38	321	13.8	150	20	AAV36003	Extended human sec
39	320	13.8	60	21	AAV00599	Human secreted pro
40	320	13.8	61	20	AAV11898	Human 5' EST secre
41	236	10.1	201	21	AAV78809	Hydrophobic domain
42	218	9.4	88	20	AAV11542	Human 5' EST secre
43	213	9.2	137	22	AAB12187	Human GPCR homolog
44	198	8.5	79	20	AAV11602	Human 5' EST secre
45	188	8.1	119	22	AAV99779	Human excretory re

ALIGNMENTS

RESULT 1

AAV57283

ID AAV57283 standard; Protein; 441 AA.

XX AC AAV57283;

XX DT 05-JUN-2000 (first entry)

XX DE Human GPCR protein (HGPRP) sequence (clone ID 1258981).

XX DE Human; G protein coupled protein receptor; HGPRP; cell proliferation;

KW neurological; immune disorder; cytostatic; anti-arteriosclerotic;

KW anti-atherosclerotic; hepatotropic; antinflammatory; virucide; leukemia;

KW immunomodulatory; anemia; asthma; gastrointestinal; anti-epileptic;

KW anti-Alzheimer's; anti-Parkinsonian; gene therapy.

XX OS Homo sapiens.

XX PN WO200015793-A2.

XX PD 23-MAR-2000.

XX PF 17-SEP-1999; 99WO-US20958.

XX PR 17-SEP-1998; 98US-0156513.

XX PA (INCY-) INCYTE PHARM INC.

XX PI Bandman O, Lal P, Tang YT, Corley NC, Guegler KJ, Gorgone GA;

XX PI Baughn MR;

XX DR WPI; 2000-271432/23.

XX DR N-PSDB; AA290521.

PT Human G protein coupled protein receptor peptides useful for the
PT prevention, diagnosis and treatment of cell proliferative, neurological
PT and immune disorders -

PS Claim 1; Page 59-60; 71pp; English.

XX The invention provides human G protein coupled protein receptor (GPCR)
CC polypeptides and polynucleotides encoding them. The polypeptides can be
CC produced by standard recombinant methodology. The polynucleotides and
CC polypeptides may be used in the prevention, treatment and diagnosis of
CC diseases associated with their inappropriate expression. Diseases that
CC can be treated are cell proliferative disorders (e.g. arteriosclerosis,
CC atherosclerosis and hepatitis), cancers (e.g. leukemia, melanomas and
CC adenocarcinoma), immune disorders (e.g. anemia, asthma and Crohn's
CC disease) and neurological disorders (e.g. epilepsy, Alzheimer's disease
CC and Parkinson's disease). The anti-GPCR antibodies may also be used as
CC diagnostic agents for detecting the presence of GPCR polypeptides in
CC samples (e.g. by enzyme linked immunosorbent assay (ELISA)). Sequences
CC AAY57283-288 represent the GPCR polypeptides.

XX Sequence 441 AA;

Query Match 100.0%; Score 2326; DB 21; Length 441;
Best Local Similarity 100.0%; Pred. No. 6.5e-241;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIHKALVMCLGLPLFLPPGAWAOGHVPYPPGCSQGLNPLLYNLCDSGAWGIVLEAVAGAG 60
Db 1 maihkalmvclglplflfpgawaoghvpypgcsqglnplynlcdrsdgaawivleavagag 60
QY 61 IVTTFVLTITILVASLPFVQDTKKRSLTGTVFFLLGTGLFCLVACVVKPDEFSTCASRR 120
Db 61 ivttfvtlilvaslpfvqdtkkrsllgtqvfllgtglfclvfacvvpkdfstcasrr 120
QY 121 FLFGVLFALFCSCLAAHVAFALNFKRNHGRGWVIFTVALLTLVEVIINTEWLIITLV 180
Db 121 flfgvlfalfcscclaaahvafalnflarknhgrgwwiftvalltlvevintewliitlv 180
QY 181 RGSCEGPGQNSAGWAVASPCAIANMDFVMAIYVMLLLGLGAFGLGAWPALCGRYKRWK 240
Db 181 rsgcegpgqngssagwawaspcaianmdfvmaaiyvmlllglgafglgawpalcgryrkwrk 240
QY 241 HGVEVLLTTATSVAIWVIMYTYGNKQNSPTWDDPTLAIALAANAWAFVLYVPIEV 300
Db 241 hgvevllttatsvaiwvivygnkqnsptwddptlaialaanawafvlyvpi 300
QY 301 SQVTKSPQSQYQGMPTTRGVGYETILKEQKQSMFVENKAFSMDPEVAAKRPVSPYSG 360
Db 301 sqvtkspqsqyqgmpttrgvgyetilkeqkqsmfvenkafsmdepvaakrpvpsyg 360
QY 361 YNGQLTSVYOPTMALMHKVPSEGAVIDIILPRATANSQVMSGANSTLRAEDMYSAQSHQ 420
Db 361 yngqltsvyoptemalmhkvpssegavidilpratansqvmgsanstlraedmysadshq 420
QY 421 AATPPKDGKNSQVFRNPVYVD 441
Db 421 aatppkdgknsqvfrnpvyvd 441

RESULT 2

AAY50932
ID AAY50932 standard; Protein; 441 AA.

XX AAY50932;

XX 10-MAR-2000 (first entry)

XX Human fetal brain cDNA clone vc26_1 derived protein #1.

XX Human; secreted protein; treatment; nutritional activity; cytokine;
KW cell proliferation; cell differentiation; hematopoiesis regulation;
KW tissue growth; activin; inhibin; chemotactic; chemokinetic; hemostatic;

thrombolytic; anti-inflammatory; invasion suppressor; tumor inhibition;
gene therapy.

OS Homo sapiens.

XX WO995721-A1.

XX 04-NOV-1999.

XX 23-APR-1999; 99WO-US08504.

XX 24-APR-1998; 98US-0082904.

XX 11-JUN-1998; 98US-0088994.

XX 12-JUN-1998; 98US-0089278.

XX 02-JUL-1998; 98US-0091647.

XX 24-AUG-1998; 98US-0097639.

XX 22-APR-1999; 99US-0097639.

XX (ALPH-) ALPHAGENE INC.

XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;

XX WPI: 2000-052801/04.

XX N-PSDB; AA243798.

XX New polynucleotides encoding secreted human proteins, derived from
PT human fetal brain, adult skin, adult brain, adult heart, adult thymus
PT and adult aorta cDNA libraries.

XX Claim 53a; Page 246-247; 282pp; English.

XX This invention describes novel human secreted proteins which are encoded
CC by polynucleotides obtained from fetal brain, adult skin, adult brain,
CC adult heart, adult thymus and adult aorta cDNA libraries. The
CC polynucleotides and proteins are predicted to have biological activities
CC which would make them suitable for treating, preventing or ameliorating
CC medical conditions in humans and animals, although no supporting data
CC is given. Suggested activities include nutritional activity, cytokine
CC and cell proliferation/differentiation activity, immune stimulating
CC (e.g. as vaccines) or suppressing activity, hematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, hemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumor
CC invasion suppressor activity, and tumor inhibition activity. The
CC polynucleotides are also stated to be useful for gene therapy.
CC AAY50905-Y50947 represent the secreted proteins described in the method
CC of the invention which are encoded by the polynucleotides represented in
CC AA243777-243808.

XX Sequence 441 AA;

Query Match 100.0%; Score 2326; DB 21; Length 441;
Best Local Similarity 100.0%; Pred. No. 6.5e-241;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIHKALVMCLGLPLFLPPGAWAOGHVPYPPGCSQGLNPLLYNLCDSGAWGIVLEAVAGAG 60
Db 1 maihkalmvclglplflfpgawaoghvpypgcsqglnplynlcdrsdgaawivleavagag 60
QY 61 IVTTFVLTITILVASLPFVQDTKKRSLTGTVFFLLGTGLFCLVACVVKPDEFSTCASRR 120
Db 61 ivttfvtlilvaslpfvqdtkkrsllgtqvfllgtglfclvfacvvpkdfstcasrr 120
QY 121 FLFGVLFALFCSCLAAHVAFALNFKRNHGRGWVIFTVALLTLVEVIINTEWLIITLV 180
Db 121 flfgvlfalfcscclaaahvafalnflarknhgrgwwiftvalltlvevintewliitlv 180
QY 181 RGSCEGPGQNSAGWAVASPCAIANMDFVMAIYVMLLLGLGAFGLGAWPALCGRYKRWK 240
Db 181 rsgcegpgqngssagwawaspcaianmdfvmaaiyvmlllglgafglgawpalcgryrkwrk 240
QY 241 HGVEVLLTTATSVAIWVIMYTYGNKQNSPTWDDPTLAIALAANAWAFVLYVPIEV 300

|||||ttatstsvaiwvmytygkqnsptwddptlaialaanaawafvfyvipev 300
Db 241 hgfvllttatstsvaiwvmytygkqnsptwddptlaialaanaawafvfyvipev 300
QY 301 SQVTKSSPEQSYQGDYMPRGVGYETILKQKQSMFVENKAFSMDEPVAARPVSPYSQ 360
Db 301 sqvtksspeqsygqdmyptrgvgyetilkqkqsgmfvenkafsmdepvaakrpvpsysq 360
QY 361 YNGQLTTSVYQPTPEMALMHKVPSEGAYDIILPRATANSQVMSGANSTLRAEDMYSAQSHQ 420
Db 361 yngqlttsvyqptemalmhkvpssegaydiilpratansqvmgsanstlraedmysaqshq 420
QY 421 AATPPKDGKNSQVFRNPYVWD 441
Db 421 aatppkdgknsqvrnpypvwd 441
RESULT 3
AAM39849
ID AAM39849 standard; Protein: 441 AA.
XX
AC AAM39849;
XX
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2994.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX 25-APR-2000; 2000US-0552317.
XX 09-JUL-2000; 2000US-0598042.
XX 19-JUL-2000; 2000US-0620312.
XX 03-AUG-2000; 2000US-0653450.
XX 14-SEP-2000; 2000US-0662191.
XX 19-OCT-2000; 2000US-0693036.
XX 29-NOV-2000; 2000US-0727344.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX N-PSDB; AA159005.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries -
XX
XX Example 4; SEQ ID NO 2994; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AA42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 441 AA;

Query Match 100.0%; Score 2326; DB 22; Length 441;
Best Local Similarity 100.0%; Pred. No. 6.5e-241;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATHKALVMCLGLPLFLPGAWAAGHVPCCSQGLNPLYYNLCDSRGANGVLEAVAGAG 60
Db 1 maIhkalvmclglplflpgawaqghvppcsgqlnplynldcrgagawivleavagag 60
QY 61 IVTTFVLTIILVASLPFVQDTKKRSLILGTQVFFLLGLTGLFVACVVKPDEFSTCASRR 120
Db 61 ivttfvtltilvaslplfvqdtkkrsllgtqvfllglfvlfcvkvkpdfstcasrr 120
QY 121 FLFGVLFAICFSCLAARHVFALNFLARKNHGPRGWIFTVALLTLVEVIINTEWLIITLV 180
Db 121 flfgvlfaicfscslaahvfalnflarknhgprgwiiftvalltlveviintewliitlv 180
QY 181 RGSCEGPGQNSAGWAVASPCAIAANMDFVMAIIYVMLLLGLGAFNGAWPALCGRYKRWK 240
Db 181 rgsgegpgqngnsagwavaspcaiamdfvmaliymllllglgafngawpalcgryrkrwk 240
QY 241 HGVEVLLTTATSVAIWVWIVMYTYGKOHNSPTWDDPTLAIALANAWAFVIFYVPEV 300
Db 241 hgvevllttatstsvaiwvmytygkqnsptwddptlaialaanaawafvifyvipev 300
QY 301 SQVTKSSPEQSYQGDYMPRGVGYETILKQKQSMFVENKAFSMDEPVAARPVSPYSQ 360
Db 301 sqvtksspeqsygqdmyptrgvgyetilkqkqsgmfvenkafsmdepvaakrpvpsysq 360
QY 361 YNGQLTTSVYQPTPEMALMHKVPSEGAYDIILPRATANSQVMSGANSTLRAEDMYSAQSHQ 420
Db 361 yngqlttsvyqptemalmhkvpssegaydiilpratansqvmgsanstlraedmysaqshq 420
QY 421 AATPPKDGKNSQVFRNPYVWD 441
Db 421 aatppkdgknsqvrnpypvwd 441

RESULT 4

AA88591
ID AA88591 standard; Protein: 441 AA.
XX
AC AA88591;
XX
DT 04-JUN-2001 (first entry)
XX
DE Human hydrophobic domain containing protein clone HP10704 #95.
XX
KW Human; hydrophobic domain; immunosuppressant; anti-HIV; neuroprotective;
KW antianemic; vulnery; antiulcer; osteopathic; anti-inflammatory;
KW cytostatic; gene therapy; autoimmune disorder; multiple sclerosis;
KW HIV infection; anaemia; burn; ulcer; osteoporosis; tumour; wound healing;
KW inflammatory bowel disease; nutritional supplement; appetite; vaccine;
KW behavioural characteristic; immune response.
XX
OS Homo sapiens.
XX
PN WO200112660-A2.
XX
XX 22-FEB-2001.
PD
XX 10-AUG-2000; 2000WO-JP05356.
XX

PR 17-AUG-1999; 99JP-0230344.
 PR 07-SEP-1999; 99JP-0252551.
 PR 01-OCT-1999; 99JP-0281132.
 PR 22-OCT-1999; 99JP-0301624.
 PR 04-NOV-1999; 99JP-0313877.
 XX (SAGA) SAGAMI CHEM RES CENT.
 PA (PROT-) PROTEGENE INC.
 XX Kato S, Kimura T;
 PI WPI; 2001-160059/16.
 XX N-PSDB; AAF94481.
 DR Human proteins with hydrophobic domains and the DNAs which encode them
 PT are useful for treating autoimmune disorders, burns and tumors and for
 PT screening novel pharmaceuticals -
 XX Claim 1; Page 413-415; 518pp; English.
 XX AAF94417 to AAF94516 encode the human proteins given in AAB88557 to
 CC AAB88606 (I) which have a hydrophobic domain. (I) have immunosuppressant,
 CC anti-Hiv, neuroprotective, antianaemic, vulnerary, antiulcer,
 CC osteopathic, anti-inflammatory and cytostatic activities, and can be
 CC used in gene therapy. (I) can be used as pharmaceuticals and as antigens
 CC to prepare antibodies. DNA and cDNA (II) encoding (I) can be used as
 CC probes for genetic diagnosis and gene sources for gene therapy or for
 CC producing (I) in large quantities. Cells containing (II) are used for
 CC the detection of ligands or receptors corresponding to membrane or
 CC secretory proteins and to screen small molecule novel pharmaceuticals.
 CC Antibodies directed to (I) can be used for the detection, quantification
 CC and purification of (I). Activities of (I) may include cytokine and cell
 CC proliferation/differentiation function, immune stimulating or suppressing
 CC activity, haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity and anti-inflammatory
 CC activity. (I) and (II) can be used to treat autoimmune disorders e.g.
 CC multiple sclerosis, HIV infections, anaemia, burns, ulcers, osteoporosis,
 CC inflammatory bowel disease and tumours. (I) and (II) can also be used for
 CC wound healing, as nutritional sources or supplements e.g. as amino acid,
 CC carbon or nitrogen source, to effect metabolism, catabolism, anabolism,
 CC processing and utilisation of dietary fat, protein, carbohydrate,
 CC vitamins and minerals, to effect behavioural characteristics, to affect
 CC appetite, and can act as antigens in vaccines to raise an immune response
 CC to the protein or another material cross-reactive with the protein.
 XX Sequence 441 AA;

Query Match 100.0%; Score 2326; DB 22; Length 441;
 Best Local Similarity 100.0%; Pred. No. 6.5e-241;
 Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATHKALVCMGLPLFLPGCAWAGHVPCCSQGLNPLYNLCDRSGAWGIVLEAVAG 60
 DB 1 mathkalmcglplflpgawagghvppcgsglnplynldcrgawgivilavag 60
 QY 61 IVTTFVLITILVASLPFQDTQVFFLLGTLGLFCLVACVVRPDEFSTCSR 120
 DB 61 ivttfvilitivaslpvfqdtqvvffllgltglfclvfacvvrpdefstcsrr 120
 QY 121 FLFGVLFATCFSCAAHFALNFLARKNHGPRGVITFVALLTLVEIVNTEWLIITLV 180
 DB 121 flfgvlfatcfsciaahfalnflarknhgrgvitfvalltlvveivntewliitlv 180
 QY 181 RGSGEQGQGNSSAGWAVASPCAIANMDFVMALYVMLLLGAFLGAWPALCGRYKRWK 240
 DB 181 rgsgegqgqgnssagwavaspcailanmdfvmaliyvmllllgaflgawpalcgryrkwrk 240
 QY 241 HGVFVLITATSVAVVWVIMVYTGKQHNSTPTWDDPTALALAAAWAFVLFVYIPEV 300
 DB 241 hgvfvlitatsvavvwwvivytygkqhnsptwddptalalaanaawafvlfvyipev 300

QY 301 SQVTKSSPEQSYQGDMPYTRGVGYETILKEQKQSGMEVENKAFSDMEPVAAKRPVSPYS 360
 DB 301 sqvtksspeqsyqgdmyptrgvgyetilkqkqsgmfvnka fmsdepvaakrpvspys 360
 QY 361 YNGQLLTSVYQPTMALMHKVPSEGAYDIILPRATANSQVMSANSTLRADMYSAQSHQ 420
 DB 361 yngqltstvqptemalmhkvpssegaydiilpratan sqvmsanstlraedmysaqshq 420
 QY 421 AATPPKDGKNSQVFRNPYVWD 441
 DB 421 aatppkdgknsqvfrnpyvwd 441
 RESULT 5
 AAB88359
 ID AAB88359 standard; Protein; 441 AA.
 XX
 AC AAB88359;
 XX
 DT 23-MAY-2001 (first entry)
 XX
 DE Human membrane or secretory protein clone PSEC0087.
 XX
 KW Human; secretory protein; membrane protein; vaccine; gene therapy;
 KW rheumatoid arthritis; diabetes.
 XX
 OS Homo sapiens.
 XX
 PN EP1067182-A2.
 XX
 PD 10-JAN-2001.
 XX
 PF 07-JUL-2000; 2000EP-0114090.
 XX
 PR 08-JUL-1999; 99JP-0194179.
 PR 11-JAN-2000; 2000JP-0118775.
 PR 02-MAY-2000; 2000JP-0183766.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
 XX WPI; 2001-093989/11.
 DR N-PSDB; AAF93786.
 XX
 PT Nucleic acids encoding secretory proteins/membrane proteins, useful in
 PT gene therapy or as candidate target molecules in drug development -
 PS
 PS Claim 1; SEQ ID 86; 609pp + CD ROM; English.
 XX
 CC This invention relates to nucleic acid sequences AAF93744 - AAF93916
 CC which encode human secretory or membrane proteins represented by
 CC AAB88317 - AAB88419. Included in the invention are primers
 CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
 CC cDNA sequences of the invention. The invention also includes methods for
 CC the production of antibodies directed against the proteins, and cDNA
 CC sequences, which can be used in vaccines. The polynucleotide sequences
 CC can be used in gene therapy. The polynucleotide sequences and the
 CC proteins they encode may be used in the prevention, treatment and
 CC diagnosis of diseases associated with inappropriate secretory
 CC protein/membrane protein expression. The nucleic acids and complementary
 CC sequences may also be used as DNA probes in diagnostic assays
 CC (e.g. polymerase chain reactions (PCR)) to detect and quantify the
 CC presence of similar nucleic acid sequences in samples. They may also be
 CC used to study the expression and function of secretory proteins/membrane
 CC polypeptides and their role in metabolism. The polypeptides may be used
 CC as antigens in the production of antibodies against them and in assays to
 CC identify modulators (agonists and antagonists) of expression and
 CC activity. The antibodies (agonists and antagonists) may also be used as therapeutic
 CC agents to down regulate expression and activity. The antibodies may also
 CC be used as diagnostic agents for detecting the presence of the
 CC polypeptides in samples (e.g. by enzyme linked immunosorbant assay
 CC (ELISA)). Examples of diseases which may be treated include rheumatoid

CC arthritis and diabetes.

XX Sequence 441 AA;

Query Match 100.0%; Score 2326; DB 22; Length 441;
Best Local Similarity 100.0%; Pred. No. 6.5e-241;
Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAIHKALVMCLGLPLFLFPGAWAQGHVPPGCSQGLNPLYNLCDRSGAWGIVLEAVAGAG 60
Db 1 maikhalmvcmclgplflfpgawagqhvpvpgcsqglnplynlcdrsgawgivleavagag 60
Qy 61 IVTTFVLTIIIVASLPFVQDTKKRSLTGTQVFFLLGTLGLFCLVFAVCWKPDPFSTCASRR 120
Db 61 ivttfvltiilvaslpfvqdtkkrrslgtqvffllgtlglfclvfcwvkpdpfstcasrr 120
Qy 121 FLFGVLFAICFSCIAAHVFALNFILARKNHGRPGWVIFTVALLTLVEIINTEWLIITLV 180
Db 121 flfgvlfaicfsciaahvfalnflarknhgrpgwviftvalltlveiintewliitlv 180
Qy 181 RGSCEGPGQNSSAGWAVASPCAIANMDFVMAIIVYVMLLLGLGAWPALCGRYKRWK 240
Db 181 rgsceggpgqnssagwavaspcaiannmdfvmaliyvmlllglgawpalcgryrkwrk 240
Qy 241 HGVFVLLTTSVAIIVWVIMVYTYGNKHNSPTWDDPTLAIALAANAWAFVIFYVPIEV 300
Db 241 hgvfllttsvaivwvwmvtygnkhnspwddptlaialaanaawafvifyvpi 300
Qy 301 SQVTKSSPEQSYQGDMPYTRGVGYETILKEQKGSFMFVENKAFSDMEPVAARKRVPSPYS 360
Db 301 sqvtksspeqsygqdmyptrgvgyetilkedqgsmfvenkafsdmepvaarkrvpsysg 360
Qy 361 YNGQLLTSVYQPTMALMHKVPSEGAYDIILPRATANSQVNGSANSILRAEDMYSAQSHQ 420
Db 361 yngqlttsvyqptemalmhkvpssegaydiilpratansqvmgsansilraedmysaqshq 420
Qy 421 AATPPKDGKNSQVFRNPYVWD 441
Db 421 aatppkdgknsqvfrnpyvwd 441

RESULT 6

AAU14166
ID AAU14166 standard; Protein; 486 AA.

AC AAU14166;

XX 24-OCT-2001 (first entry)

XX Human novel protein #37.

XX Human: novel protein; Antianemic; osteopathic; antinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnerary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.

XX Homo sapiens.

XX WO200155437-A2.

PN 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US02623.

PF 25-JAN-2000; 2000US-0491404.

PR (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

PI

XX

DR WPI: 2001-451939/48.

DR N-PSDB; AAS22471.

XX

PT Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage -
XX Example 4; Page 545-546; 894pp; English.

XX

CC The invention relates to polynucleotides encoding novel human
CC proteins or their active domains. The polypeptides, polynucleotides and
CC antibodies raised against the polypeptides are used in a method of
CC treatment of a mammal and prevention of disorders caused by the aberrant
CC protein expression or activity. The polypeptides can be used as
CC molecular weight markers, food supplements, and in antibody production.
CC The polypeptides are used to identify compounds which bind to the
CC polypeptides. Polynucleotides of the invention are used as probes and
CC primers, for sequencing, for chromosome or gene mapping, in the
CC production of recombinant proteins, and in generating anti-sense DNA or
CC RNA and in gene therapy. Polypeptides of the invention can be used to
CC target drugs to a tumour, in assays to determine biological activity, to
CC raise antibodies/ elicit an immune response, to determine quantitative
CC protein levels, as tissue markers, and to isolate receptors or ligands.
CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue, wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma,
CC graft-versus-host disease, eczema, haemophilia, thrombosis,
CC anti-inflammatory diseases, nervous system disorders, and infection.
CC The present sequence represents a protein of the invention.

XX Sequence 486 AA;

SQ

Query Match 100.0%; Score 2326; DB 22; Length 486;

Best Local Similarity 100.0%; Pred. No. 7.6e-241;

Matches 441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAIHKALVMCLGLPLFLFPGAWAQGHVPPGCSQGLNPLYNLCDRSGAWGIVLEAVAGAG 60
Db 46 maikhalmvcmclgplflfpgawagqhvpvpgcsqglnplynlcdrsgawgivleavag 105
Qy 61 IVTTFVLTIIIVASLPFVQDTKKRSLTGTQVFFLLGTLGLFCLVFAVCWKPDPFSTCASRR 120
Db 106 ivttfvltiilvaslpfvqdtkkrrslgtqvffllgtlglfclvfcwvkpdpfstcasrr 165
Qy 121 FLFGVLFAICFSCIAAHVFALNFILARKNHGRPGWVIFTVALLTLVEIINTEWLIITLV 180
Db 166 flfgvlfaicfsciaahvfalnflarknhgrpgwviftvalltlveiintewliitlv 225
Qy 181 RGSCEGPGQNSSAGWAVASPCAIANMDFVMAIIVYVMLLLGLGAWPALCGRYKRWK 240
Db 226 rgsceggpgqnssagwavaspcaiannmdfvmaliyvmlllglgawpalcgryrkwrk 285
Qy 241 HGVFVLLTTSVAIIVWVIMVYTYGNKHNSPTWDDPTLAIALAANAWAFVIFYVPIEV 300
Db 286 hgvfllttsvaivwvwmvtygnkhnspwddptlaialaanaawafvifyvpi 345
Qy 301 SQVTKSSPEQSYQGDMPYTRGVGYETILKEQKGSFMFVENKAFSDMEPVAARKRVPSPYS 360
Db 346 sqvtksspeqsygqdmyptrgvgyetilkedqgsmfvenkafsdmepvaarkrvpsysg 405
Qy 361 YNGQLLTSVYQPTMALMHKVPSEGAYDIILPRATANSQVNGSANSILRAEDMYSAQSHQ 420
Db 406 yngqlttsvyqptemalmhkvpssegaydiilpratansqvmgsansilraedmysaqshq 465
Qy 421 AATPPKDGKNSQVFRNPYVWD 441
Db 466 aatppkdgknsqvfrnpyvwd 486

```
RESULT 7
AA49156
ID AA49156 standard; Protein; 441 AA.
XX
AC AA49156;
XX
DT 17-JAN-2000 (first entry)
XX
DE G-protein coupled receptor AXOR4 amino acid sequence #1.
XX
KW AXOR4; G-protein coupled receptor; RAIG1; pain; cancer; diabetes;
KW obesity; anorexia; bulimia; asthma; stroke; ulcer; allergy; infertility;
KW neurological disorder; schizophrenia.
XX
OS Homo sapiens.
XX
PN W09953054-A1.
XX
PD 21-OCT-1999.
XX
PF 07-APR-1999; 99WO-GB01067.
XX
PR 08-APR-1998; 98GB-0007723.
XX
PA (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI Medhurst A, Michalovich D, Pangalos M, Hill J;
XX
DR WPI; 1999-620425/53.
XX
DR N-PSDB; AAZ32501.
XX
PT New AXOR4 G-protein coupled receptor polypeptides and related nucleic
PT acid, for treating, preventing or diagnosing e.g. cancer
XX
PS Claim 2; Page 38-39; Sipp; English.
XX
CC This is the G-protein coupled receptor AXOR4, amino acid sequence #1. The
CC AXOR4 polypeptide has homology with human putative G-protein coupled
CC receptor RAIG1. The AXOR4 polypeptide and polynucleotide sequences can be
CC used in the production of antibodies specific for AXOR4. The sequences
CC may also be used to screen for AXOR4 agonists or antagonists.
CC Alternatively the effect of a candidate agonist or antagonist compound on
CC the production of mRNA encoding AXOR4 may be detected using an ELISA
CC assay. Diseases or conditions arising from altered expression or activity
CC of AXOR4 may be diagnosed by detecting the AXOR4 protein in a sample from
CC a patient or detecting a mutation in the AXOR4 nucleotide sequence in the
CC genome of a patient. These diseases or conditions include pain, cancer,
CC diabetes, obesity, anorexia, bulimia, asthma, hypo- or hypertension,
CC stroke, ulcers, allergy, benign prostatic hypertrophy, migraine,
CC epilepsy, vomiting, psychosis, infertility, and neurological disorders
CC (e.g. anxiety, schizophrenia, depression, delirium, dementia,
CC neurodegeneration or severe mental retardation). Detection of the AXOR4
CC nucleotide sequence may be used for chromosome mapping and tissue
CC localization. The polynucleotide and polypeptide sequences may be
CC administered to patients as vaccines or as part of a gene therapy regime
CC respectively, to treat the diseases and conditions listed above.
XX
SQ Sequence 441 AA;
```

```
Query Match 100.0%; Score 2325; DB 20; Length 441;
Best Local Similarity 99.88; Pred. No. 8.4e-241;
Matches 440; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAIHKALVMCLGLPLFLFPGWAQHVPVPGCSQGLNPLYNLCDRSGAWGIVLEAVAGAG 60
Dy 1 maihkalmvclglplflfpgwaaghvppgcsqglnplynlncldrsgawgivleavag 60
Qy 61 IVTTFVLTILVASLPFVQDQTKRSLGTQVFFLLGTGLGLFCLVACVVKPDFSTCSARR 120
Dy 61 ivttftvltilvaslpfvqdtkrslgtqvffllgtlglfclvfacvvpkpdfstcsarr 120
```

```
Qy 121 FLFGVLFALCFSCLAHVFALNFKRKHGPRGWVFTVALLTLVEIINTWLIITLV 180
Dy 121 flfgvlfalcfsclaahvfalnfnlarknhgprgwwvftvalltlveintewliitlv 180
Qy 181 RGSSEGGPQGNSSAGWAVASPCAIANMDFVMALYVVMILLGAFLGAWPALCGRYKRWK 240
Dy 181 rgsgseggpqgnssagwavaspcavnmfdfmallyvmilllgaflgawpalcgrkwrk 240
Qy 241 HGVFVLLTATSAIVWVIMVYTYGNKHNSPTWDDPTLAIALAANAWAFVLYVIVPEV 300
Dy 241 hgvfvlltatsvaivwvviwmytygnkhnsptwddptlaialaanaawatvlyvipev 300
Qy 301 SOVTKSPQEQYQGDMPYTRGVYETILKEQGMFVENKAFSMDPEVAAKRPVSPYSG 360
Dy 301 sovtkspeqyqgdmpytrgvgyetllkqgsmfvenkafsmdpevvaakrpvspsyg 360
Qy 361 YNGQLTTSVYQPTMALMHKVPSEGYDIILPRATANSQVMSANSTLRAEDMYSQASHQ 420
Dy 361 yngqlttsvyqptemalmhkvpssegaydiilpratansqvmgsanstlraedmysaqshq 420
Qy 421 AATPPKDGKNSQVFRNPYVWD 441
Dy 421 aatppkdgknsqvfrnpyvwd 441

RESULT 8
AA41635
ID AA41635 standard; Protein; 460 AA.
XX
AC AA41635;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6566.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN W0200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US34263.
XX
PR 21-JAN-2000; 2000US-0488725.
XX
PR 25-APR-2000; 2000US-0552317.
XX
PR 09-JUL-2000; 2000US-0598042.
XX
PR 19-JUL-2000; 2000US-0620312.
XX
PR 03-AUG-2000; 2000US-0653450.
XX
PR 14-SEP-2000; 2000US-0662191.
XX
PR 19-OCT-2000; 2000US-0693036.
XX
PR 29-NOV-2000; 2000US-0727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
XX
DR N-PSDB; AAI60791.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT such as central nervous system injuries
XX
PS Example 2; SEQ ID NO 6566; 10078pp; English.
```

xx The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAW42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
xx
SQ Sequence 460 AA;

Query Match 99.7%; Score 2320; DB 22; Length 460;
Best Local Similarity 99.8%; Pred. No. 3.1e-240;
Matches 440; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAIHKALVCMGLPLFLFPAGAWAGHVPVPGCSQGLNPLYNLCDSGAWGIVLEAVAGAG 60
Db 20 mahlkalvmcglplflfpgawagghvppgcsqglnplynldcsgawgivleavagag 79

QY 61 IVTTFVLTIIIVASLPFVQDTKKRSLGTQVFFLLGLTGLFCLVFCACVVPDFSTCASRR 120
Db 80 ivttfvltiilvaslpfvqdtkkrrllgtqvfllgtlglfclvfcaevkpdfstcasrr 139

QY 121 FLFGVLFALCFSCLAHVFALNFALRNKHGPRGWVFTVALLTLVEVINTEWLIITLV 180
Db 140 flfgvlfalcfscslaahvfalnflarhknhgrgwwvftvalltlvevintewliitlv 199

QY 181 RGSSEGGPOGNSAGWAVASPCAIANMDFVMAIIVYMLLLGAFLGAWPALCGRYKRWK 240
Db 200 rgssegppgnsagwawaspcaianmdfvmaliivymlllgaflgawpalcgryrkwrk 259

QY 241 HGVFVLTATSAIVAIWVIMVITYGNKHNSPTWDDPTLAIALAANAFAVLFYVPEV 300
Db 260 hgvfvlattatsaivaiwvivygnkhnspwtddptlaialaanaafvlyfypev 319

QY 301 SQVTKSPQSYQDMPYTRCGVETILKEQKGSQSMFVENKAFSDPEVAAKRPVSPYSG 360
Db 320 sqvtkspqsyqgdmyptrcgvyetilkqkqsgsmfvenkafsdpevvaakrpvpsyg 379

QY 361 YNGOLLTSVYQPTMALMHKVPSEGYDIIILPRATANSQVNGSANSSTLRAEDMYSAQSHQ 420
Db 380 yngolltsvyqptemalmhkvpsegaydiiilpratansqvnsgsantlraedmysaqshq 439

QY 421 AATPPKDGKNSQVRNPYVMD 441
Db 440 aatppkdgknsqvrnpyvmd 460

RESULT 9
ID AAY44273
XX AAY44273 standard; Protein; 479 AA.
XX
AC AAY44273;
XX
DT 28-FEB-2000 (first entry)
XX
DE Human Metabotropic Glutamate Receptor-like protein, MGRcm.
XX
KW Metabotropic Glutamate Receptor-like protein; MGLuR; MGRcm; G-protein;
KW 7-transmembrane; GABA receptor; diagnostic marker; pharmaceutical agent;
KW central and peripheral nervous system; cardiac disorder; treatment;
KW urologic disorder; gastrointestinal disorder; diagnosis; seizure; anoxia;
KW epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;

KW antibody; inhibitor; screening.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 20..40 /label= Extracellular_domain
FT /note= "Relatively short"
FT Domain 40..300 /note= "Group of 7 transmembrane segments"
FT Domain 301..479 /label= C-terminal_cytosolic_domain
FT /note= "Relatively long"
PN WO9960121-A1.
XX
XX 25-NOV-1999.
XX
XX 19-MAY-1999; 99WO-IL00265.
XX
XX 19-MAY-1998; 98US-0085973.
XX (COMP-) COMPUGEN LTD.
XX Mintz L, Savitsky K, Toporik A;
PI
XX
DR WPI: 2000-086596/07.
DR N-PSDB; AA29292.
XX
PT New metabotropic glutamate receptor-like protein, useful for, e.g.
PT diagnosing neurological diseases
XX
PS Claim 1; Fig 1B; 84pp; English.
XX
CC The present amino acid sequence is the human metabotropic
CC glutamate receptor like protein (MGLuR-like receptor protein), MGRcm. It
CC is a novel member of the metabotropic glutamate/GABA receptor class of
CC 7-transmembrane G-protein coupled receptors showing sequence similarity.
CC It is responsible for activation of intracellular G-proteins. The MGRcm
CC receptor protein and encoding nucleic acid sequence, are useful as
CC diagnostic markers and pharmaceutical agents for various diseases of the
CC central and peripheral nervous system, as well as cardiac, urologic and
CC gastrointestinal disorders. It can be used for the diagnosis and
CC treatment of neurological and neurodegenerative disorders, such as
CC seizures, epilepsy, anoxia, Alzheimer's disease, Huntington's disease,
CC Parkinson's disease, etc. The antibody specific against MGRcm has
CC diagnostic and therapeutic applications against these disease states.
CC This receptor protein can also be employed for screening of potential
CC peptide or small molecule inhibitors of the relevant receptor-ligand
CC interaction.
XX
SQ Sequence 479 AA;

Query Match 95.7%; Score 2226; DB 21; Length 479;
Best Local Similarity 100.0%; Pred. No. 4.2e-230;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAIHKALVCMGLPLFLFPAGAWAGHVPVPGCSQGLNPLYNLCDSGAWGIVLEAVAGAG 60
Db 46 mahlkalvmcglplflfpgawagghvppgcsqglnplynldcsgawgivleavagag 105

QY 61 IVTTFVLTIIIVASLPFVQDTKKRSLGTQVFFLLGLTGLFCLVFCACVVPDFSTCASRR 120
Db 106 ivttfvltiilvaslpfvqdtkkrrllgtqvfllgtlglfclvfcaevkpdfstcasrr 165

QY 121 FLFGVLFALCFSCLAHVFALNFALRNKHGPRGWVFTVALLTLVEVINTEWLIITLV 180
Db 166 flfgvlfalcfscslaahvfalnflarhknhgrgwwvftvalltlvevintewliitlv 225

QY 181 RGSSEGGPOGNSAGWAVASPCAIANMDFVMAIIVYMLLLGAFLGAWPALCGRYKRWK 240
Db 226 rgssegppgnsagwawaspcaianmdfvmaliivymlllgaflgawpalcgryrkwrk 285

CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma,
 CC graft-versus-host disease, eczema, haemophilia, thrombosis,
 CC anti-inflammatory diseases, nervous system disorders, and infection.
 CC The present sequence represents a protein of the invention.
 XX
 SQ Sequence 231 AA;

Query Match 52.5%; Score 1221; DB 22; Length 231;
 Best Local Similarity 100.0%; Pred. No. 1.4e-122;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 211 MALIYVMLLLGAFGLNPAALCGRYKRWKRGVFLVLTATTSVAIVWVIMVYTGKQH 270
 DB 1 MALIYVMLLLGAFGLNPAALCGRYKRWKRGVFLVLTATTSVAIVWVIMVYTGKQH 60
 QY 271 NSPTWDDPTLAIALANAWAFVIVPEVSQVTKSSPEQSYQGDMPYTRGVGYETILKE 330
 DB 61 NSPTWDDPTLAIALANAWAFVIVPEVSQVTKSSPEQSYQGDMPYTRGVGYETILKE 120
 QY 331 QKGSMFVENKAFSMDEPVAARKRPVSPYSGYNGQLTSVQPTMALMHKVPSEGAYDII 390
 DB 121 QKGSMFVENKAFSMDEPVAARKRPVSPYSGYNGQLTSVQPTMALMHKVPSEGAYDII 180
 QY 391 LPRATANSQVNGSANSFLRADMYSAQSHQAATPPKDGKNSQVFRNPYVWD 441
 DB 181 LPRATANSQVNGSANSFLRADMYSAQSHQAATPPKDGKNSQVFRNPYVWD 231

RESULT 13
 AAY44274
 ID AAY44274 standard; Protein; 296 AA.

AC AAY44274;
 XX
 XX 28-FEB-2000 (first entry)
 XX
 DE Mouse Metabotropic Glutamate Receptor-like protein, MGRcm.
 XX
 KW Metabotropic Glutamate Receptor-like protein; MGLur; MGRcm; G-protein;
 KW 7-transmembrane; GABA receptor; diagnostic marker; pharmaceutical agent;
 KW central and peripheral nervous system; cardiac disorder; treatment;
 KW urologic disorder; gastrointestinal disorder; diagnosis; seizure; anoxia;
 KW epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;
 KW antibody; inhibitor; screening; mouse; homologue.
 XX
 OS Mus musculus.

XX Key Location/Qualifiers
 FT Protein 1..296
 FT /label= Mouse_MGRcm_protein
 FT /note= "Homolog of human MGRcm"
 FT Misc-difference 37
 FT /note= "Corresponds to CYA codon"
 FT Misc-difference 148
 FT /note= "Corresponds to TWV codon"
 FT Misc-difference 155
 FT /note= "Corresponds to WGC codon"

XX WO960121-A1.
 XX
 XX 25-NOV-1999.
 XX
 XX 19-MAY-1999; 99WO-IL00265.
 XX
 XX 19-MAY-1998; 98US-0085973.
 XX
 XX (COMP-) COMPUGEN LTD.

PI Mintz L, Savitsky K, Toporik A;
 XX
 DR WPI: 2000-086596/07.
 DR N-PSDB; AAZ29293.
 XX
 PT New metabotropic glutamate receptor-like protein, useful for, e.g.
 PT diagnosing neurological diseases -
 XX
 PS Claim 1; Fig 2B; 84pp; English.
 XX
 CC The present amino acid sequence is the mouse metabotropic
 CC glutamate receptor like protein (MGLur-like receptor protein), MGRcm. It
 CC is a novel member of the metabotropic glutamate/GABA receptor class of
 CC 7-transmembrane G-protein coupled receptors showing sequence similarity.
 CC It is homologous to the human MGRcm sequence.
 CC It is responsible for activation of intracellular G-proteins. The MGRcm
 CC receptor protein and encoding nucleic acid sequence, are useful as
 CC diagnostic markers and pharmaceutical agents for various diseases of the
 CC central and peripheral nervous system, as well as cardiac, urologic and
 CC gastrointestinal disorders. It can be used for the diagnosis and
 CC treatment of neurological and neurodegenerative disorders, such as
 CC seizures, epilepsy, anoxia, Alzheimer's disease, Huntington's disease,
 CC Parkinson's disease, etc. The antibody specific against MGRcm has
 CC diagnostic and therapeutic applications against these disease states.
 CC This receptor protein can also be employed for screening of potential
 CC peptide or small molecule inhibitors of the relevant receptor-ligand
 CC interaction.
 XX
 SQ Sequence 296 AA;

Query Match 51.6%; Score 1200.5; DB 21; Length 296;
 Best Local Similarity 84.9%; Pred. No. 3.4e-120;
 Matches 230; Conservative 11; Mismatches 29; Indels 1; Gaps 1;

QY 1 MATHKALVMCLGULPLFLPGAWAQGHVPPGCSQGLNPLYYNLCDRSGAWGIVLEAVAG 60
 DB 27 mathrtlmcxgplf-fpgalaqnahppgcspldplynalcdrsgawgivleavag 85
 QY 61 IVTTFVLTITILVASLPFVQDTKRSLGTVQVFFLLGLFLCFLVACVVKPDEFSTCASRR 120
 DB 86 iitffvltitilvaslpfvqdtkrslgtqvfllgtlglfcflvacvkvkpfstcasrr 145
 QY 121 FLFGVLFPAICFCLAHVFAINFLARKNHGPRGVVITVALLTLVEVIINTEWLIITLV 180
 DB 146 flxgvlfaixfscivahvlsinftrknhgrgwwviftvalltlveviintewliitlv 205
 QY 181 RGSGEQPGQNSSAGWAVASPCAIANMDFVMAIYVMLLLGAFGLNPAALCGRYKRWK 240
 DB 206 pggsgvqplanvsadstmtspcalanmdfmaliyvmlllltalfgawptlccrfrwrk 265
 QY 241 HGVFVLLTATTSVAIVWVIMVYTGKQHN 271
 DB 266 hgvfvllttvtsiaivwvwmvmytygneqhn 296

RESULT 14
 AAY50933
 ID AAY50933 standard; Protein; 302 AA.
 XX
 AC AAY50933;
 XX
 XX 10-MAR-2000 (first entry)
 XX
 DE Human fetal brain cDNA clone vc26_1 derived protein #2.
 XX
 KW Human; secreted protein; treatment; nutritional activity; cytokine;
 KW cell proliferation; cell differentiation; hematopoiesis regulation;
 KW tissue growth; activin; inhibitor; chemotactic; chemokinetic; hemostatic;
 KW thrombolytic; anti-inflammatory; invasion suppressor; tumor inhibition;
 KW gene therapy.
 XX
 OS Homo sapiens.


```

XX  W09955721-A1.
XX  04-NOV-1999.
XX  23-APR-1999; 99WO-US08504.
XX  24-APR-1998; 98US-0082904.
XX  11-JUN-1998; 98US-0088994.
XX  12-JUN-1998; 98US-0089278.
XX  02-JUL-1998; 98US-0091647.
XX  24-AUG-1998; 98US-0097639.
XX  22-APR-1999; 99US-0097639.
XX  (ALPH-) ALPHAGENE INC.
XX  Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
XX  WPI; 2000-052801/04.
XX  N-PSDB; AA243798.
XX  New polynucleotides encoding secreted human proteins, derived from
XX  human fetal brain, adult skin, adult brain, adult heart, adult thymus
XX  and adult aorta cDNA libraries.
XX  Disclosure; Page 274-275; 282pp; English.
XX  This invention describes novel human secreted proteins which are encoded
XX  by polynucleotides obtained from fetal brain, adult skin, adult brain,
XX  adult heart, adult thymus and adult aorta cDNA libraries. The
XX  polynucleotides and proteins are predicted to have biological activities
XX  which would make them suitable for treating, preventing or ameliorating
XX  medical conditions in humans and animals, although no supporting data
XX  is given. Suggested activities include nutritional activity, cytokine
XX  and cell proliferation/differentiation activity, immune stimulating
XX  (e.g. as vaccines) or suppressing activity, hematopoiesis regulating
XX  activity, tissue growth activity, activin/inhibin activity,
XX  chemotactic/chemokinetic activity, hemostatic and thrombolytic activity,
XX  receptor/ligand activity, anti-inflammatory activity, cadherin/tumor
XX  invasion suppressor activity, and tumor inhibition activity. The
XX  polynucleotides are also stated to be useful for gene therapy.
XX  AA30905-Y30947 represent the secreted proteins described in the method
XX  of the invention which are encoded by the polynucleotides represented in
XX  AA243777-243808.
XX  Sequence 302 AA;

Query Match 43.9%; Score 1020.5; DB 21; Length 302;
Best Local Similarity 66.6%; Pred. No. 7.8e-101;
Matches 209; Conservative 10; Mismatches 30; Indels 65; Gaps 5;

QY 1 MATHKALVMCLGLPLFLFPGAWAQGHVPPGCSQGLNPLYNLCDRSGANGIVLEAVAG 60
Db 1 mathkalvmclglplflfgawaqghvppgcsqglnplynldcrgsgangivleavag 60
QY 61 IVTFVTLITLVASLPVQDTKRSLTGTVFLLGLFGLVFACVKVPDFSTCASRR 120
Db 61 ivtftvtilvaslpfvqdtkrslrgtqvfllglfglvfcvkvkpdfstcasrr 120
QY 121 FLFGVLFAICFSCLAARFALNFALRNHGRGWVFTVALLTLVEVIINTEWLIITLV 180
Db 121 flfgvlfaicfscslaahvafalnfnarknhrgrgwvftvalltlveviintewliitlv 180
QY 181 RGSGBGGPQGNSSA-----GWAVASCAITANMDFVM 211
Db 181 rsgsgbggpqgnssaaagppvpptwlswhstscocccwpswpgpvcvaats----- 236
QY 212 ALIYVMLLLLGAFGLAWPALCGRYK-----RWRKHGVFVLLTTATSVAILVWVIMYTY 265
Db 237 -----agvsmg---sicssppqpplypgwgcgssclitatssttvpvgm----- 276
QY 266 GNKHNSPTWDDPT 279

```

```

Db 277 -----tpwpsps 284
RESULT 15
AA32141
ID AAY32141 standard; Protein; 403 AA.
XX AAY32141;
XX 01-FEB-2000 (first entry)
XX Human G-protein coupled receptor AXOR3.
XX AXOR3; G-protein coupled receptor; human; pain; cancer;
XX diabetes; obesity; anorexia; bulimia; asthma; hypotension;
XX hypertension; stroke; ulcer; allergy; benign prostatic hypertrophy;
XX migraine; epilepsy; vomiting; psychosis; infertility;
XX neurological disorder; anxiety; schizophrenia; manic depression;
XX depression; delirium; dementia; neurodegenerative disease;
XX mental retardation; therapy; diagnosis.
XX Homo sapiens.
XX W09952944-A1.
XX 21-OCT-1999.
XX 06-APR-1999; 99WO-GB01038.
XX 08-APR-1998; 98GB-0007722.
XX (SMIK ) SMITHKLINE BEECHAM PLC.
XX Medhurst A, Michalovich D, Pangalos M, Hill J;
XX WPI; 1999-633823/54.
XX N-PSDB; AA220297.
XX New isolated G-protein coupled receptor AXOR3 polypeptides, used to
XX develop products for treating, e.g. pain, cancers, neurological and
XX neurodegenerative disorders.
XX Claim 1; Page 40-41; 51pp; English.
XX The present sequence represents AXOR3, a novel human G-protein
XX coupled receptor that shows homology to the human G-protein coupled
XX receptor RAIG1. The invention relates to AXOR3 polypeptides and
XX polynucleotides, recombinant materials and methods for their
XX production. It also relates to methods for using such polypeptides
XX and polynucleotides for treatment of pain, cancers, diabetes,
XX obesity, anorexia, bulimia, asthma, hypotension, hypertension,
XX stroke, ulcers, allergies, benign prostatic hypertrophy, migraine,
XX epilepsy, vomiting, psychosis, infertility, neurological disorders
XX including anxiety, schizophrenia, manic depression, depression,
XX delirium, dementia, neurodegenerative diseases and severe mental
XX retardation. Also disclosed are methods for identifying agonists
XX and antagonists/inhibitors, and treating conditions associated with
XX AXOR3 imbalance with the identified compounds. The invention also
XX relates to diagnostic assays for detecting diseases associated with
XX inappropriate AXOR3 activity or levels.
XX Sequence 403 AA;

Query Match 31.5%; Score 733; DB 20; Length 403;
Best Local Similarity 39.4%; Pred. No. 9.8e-70;
Matches 164; Conservative 60; Mismatches 114; Indels 78; Gaps 9;

QY 1 MATHKALVMCLGLPLFLFPG-ANAQGHVPPGCSQGLNPLYNLCDRSGANGIVLEAVAGA 59
Db 9 mrqhgvltfil---lfvitsvasenastrgcgldilpqvsvicdldaiwgivveavaga 65

```

```
QY 60 GIVTFVLTITILVASLPFVODTKRSLGTVQVFELLGTLCLFCLVFAVVKPDPFSTCASR 119
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
66 galitllmlllvrpfikekkspgvghflflgtlglfllgtlgtfaiiqedeticsvr 125
QY 120 RFLGVLFAICFCLAAHVAFALNPLARKNHGPRGWVFTVALLTLVEVINTEWLIIITL 179
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
126 rflwgvlfaicfclscsqawrvrlvrhgtgpagwqlvgialclmlvqvliavewlvltv 185
QY 180 VRSGEGGPGQGNSSAGWAVASPCAIAANMDFVMAIYYVMLLLIGAFGLGAWPALCGRYKRW 239
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
186 lrdt-----tpacayepmdfvmaliydmvllvvtlglalftlclgkfrwk 230
QY 240 KHGVFVLLTTATSVAIWVWVIMVITYGN-KOHNSPTWDDPTLAIALAANAFAVLFYVIP 298
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
231 lngaflilitafslwiwvmtmylfgnvklqgdawndptlaitlaasgwvfifhaip 290
QY 299 EV-----SOVTKSSP---EOSYQGDMPYTRGVGYETILKEQKGQSMFVENKA 342
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
291 eihtllpalqentpnyfdtsqprmrtafeedvqlpra-----ymenka 335
QY 343 FMSDEPVAA-----KRPVSPSYSGYNGQLLTSVYQPTMALM 378
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
336 fsmdehnaalrtagfngsglgrpsgslgrpsapfr-----snvyqptemavv 384
```

Search completed: September 16, 2002, 09:32:18
Job time: 327 sec